DOCUMENT RESUME

ED 160 109 IR 006 458

AUTHOR Markey, Karen; Atherton, Pauline

TITLE ONTAP: Online Training and Practice Manual for ERIC Data Base Searchers.

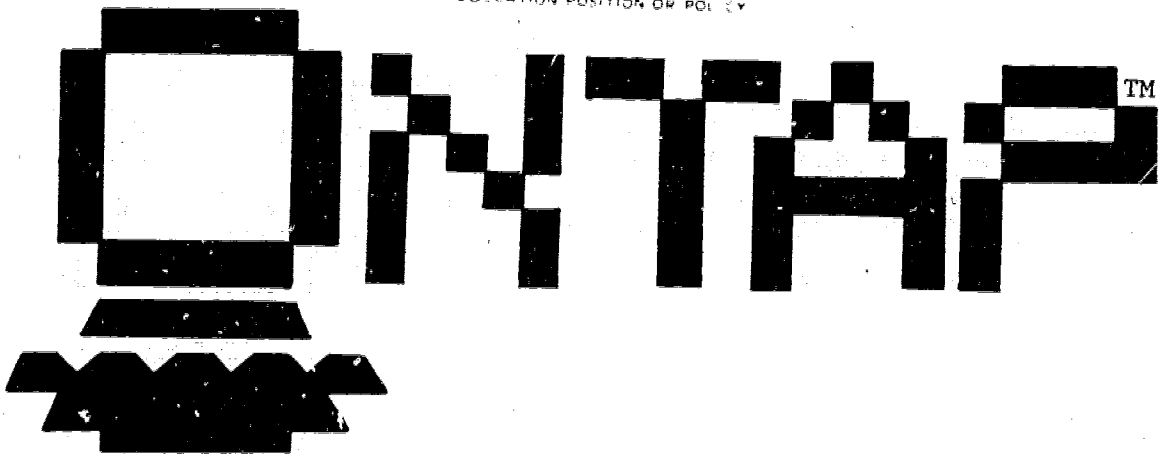
INSTITUTION Syracuse Univ., N.Y. ERIC Clearinghouse on Information Resources.

SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Jun 78
GRANT NIE-R-76-0027
NOTE 182p.; Parts of document marginally legible
AVAILABLE FROM Syracuse University Printing Services, 125 College Place, Syracuse, New York 13210 (IR-24; $6.00 plus $0.60 postage and handling)

EDRS PRICE MF-$0.83 HC-$10.03 Plus Postage.
DESCRIPTORS *Data Bases; Improvement Programs; Library Education; *Manuals; *Models; On Line Systems; Relevance (Information Retrieval); *Search Strategies; *Training
IDENTIFIERS DIALOG; *ERIC; Information Analysis Products; Information Specialists

ABSTRACT

This self-improvement manual developed for experienced online searchers presents and discusses several alternative search strategies and their appropriate use to achieve the desired outcome. The manual is divided into two parts: (1) an eight step model of the total search process emphasizing how the decision of the information professional affects retrieval results and search objectives at each step of the search process; and (2) introduction to DIALOG's ONTAP file and self-improvement exercises with search formulations and explanations selected from the questions stored in the ONTAP file. Each exercise falls within the scope of interest of one of the 16 ERIC Clearinghouses. The ONTAP file is a subset of the ERIC data base consisting of RIE and CIJE citations for the 1975 ERIC file and 29 simple, moderate, or difficult search questions with answer sets. It was developed to allow searchers to evaluate their output by computing recall and precision scores of their results with the results of a "perfect search" or answer set. The appendices include 13 search save formulations of common search facets for "population level" and "tools" which can be accessed and exchanged between DIALOG workspaces by using the procedures that are outlined, and a comprehensive guide to DIALOG's ERIC data bases. (JPF)



ONLINE TRAINING AND PRACTICE MANUAL FOR ERIC DATA BASE SEARCHERS

BY KAREN MARKEY AND PAULINE ATHERTON

BASED ON THE EARLIER UNPUBLISHED WORK OF
CHARLES P. BOURNE, BARBARA ANDERSON,
AND JO ROBINSON

ERIC CLEARINGHOUSE ON INFORMATION RESOURCES
SYRACUSE UNIVERSITY
JUNE 1978

Karen Markey is a Research Associate in the School of Information Studies, Syracuse University, Syracuse, New York.

Pauline Atherton is a Professor in the School of Information Studies, Syracuse University, Syracuse, New York.

The material in this publication was prepared pursuant to a contract with the National Institute of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their judgement in professional and technical matters. Prior to publication, this document was submitted to Lockheed Information Systems, Palo Alto, California and to independent online searchers for critical review and determination of professional competence. This publication has met such standards. Points of view or opinions, however, do not necessarily represent the official view or opinions of either the reviewers or the National Institute of Education.

Prepared under Grant No. NIE-R-76-0027.

TABLE OF CONTENTS

PREFACE

As more and more online searchers gain experience, there appears to be a need for online and offline aids to support searchers through the "improver" stage of training and practice. Few examples serving this purpose can be found in the literature, but we learned of a project in this area in early 1977. At the suggestion of, and with the encouragement and support of Oscar Firschein at Lockheed Information Systems, Charles Bourne and his colleagues at the Institute for Library Research (ILR) at the University of California, Berkeley, were developing a special ERIC/DIALOG file for online training and practice. This file was being specially designed to allow online searchers to compare their search results with the results of a "perfect search." The Berkeley group developed "answer sets" for some 20 search requests by exhaustively searching a file of ERIC 1975 citations for relevant hits. The composite results of all their search strategies for each question became the "perfect answer set." For each search topic, they offered several representative formulations to meet different search objectives (e.g., high recall, high precision, or low cost).

Oscar Firschein arranged to have 1975 ERIC accessions, and the 29 search questions with their answer sets loaded on the DIALOG system. He developed and implemented the /EVAL command on the DIALOG system. This automatically computes a recall and precision score for any searcher's effort on the ONTAP™ questions. by comparing the searcher's results with the "perfect answer set."

A draft report was prepared by the ILR group to document their work, but it could not be published before they left ILR. In late 1977 at ERIC/IR, we decided to work from their draft, with their permission, and develop an "improver's" manual for searching the ONTAP data base online.

The purpose of this publication is neither to train beginning searchers nor to provide an orientation to computer-based searching of ERIC. Publications

already perform these functions quite admirably.[1]

Our objective has been to develop the Bourne, Robinson and Anderson draft report into a manual that the experienced searcher can use for self-improvement. In the first section, a model of the total search process is discussed in order to emphasize how the decisions of the information professional affect retrieval results and search objectives at each step of the process. The second section begins with an introduction to DIALOG's ONTAP file which is followed by "self improvement" exercises selected from the questions stored in the ONTAP file. The first exercise provides an in-depth explanation of the ONTAP file as directions from the ONTAP file, the EVAL capability, and selecting self-improvement exercises are discussed in detail. Within a self-learning environment, the searcher is encouraged to analyze and compare his/her search objectives and output with that of other searchers as one proceeds with the exercises.

To make this publication as useful as possible to ERIC Clearinghouse personnel, we have identified the self-improvement exercises by ERIC Clearing-

---

[1] Bourne, Charles P. et al. Analysis of ERIC On-Line File Searching. Procedures and Guidelines for Searching. Final Report. Berkeley, CA: University of California, Institute of Library Research, and Palo Alto, CA: Lockheed Research Laboratories, Nov. 1974. (ED 101 757)

Bourne, Charles P. DIALOG Lab Workbook; training exercises for the Lockheed DIALOG Information Retrieval Service. Berkeley, CA: University of California, Institute of Library Research, Oct. 1976. (ED 136 751)

Brandhorst, W. T., and Sharon Jewell. Search Strategy Tutorial; Searcher's Kit. Washington, DC: National Institute of Education, Oct. 1973. (ED 082 763)

Timbie, Michele, and Don H. Coombs. An Interactive Information Retrieval System; case studies on the use of DIALOG to search the ERIC document file. Stanford University, ERIC Clearinghouse on Educational Media and Technology, Dec. 1969. (ED 034 431)

Yarborough, Judith. How to Prepare for a Computer Search of ERIC: a non-technical approach. 44p. Sept. 1975. (ED 110 096, $.83 microfiche, $2.06 photocopy, plus postage, available from EDRS).

Yarborough, Judith. "A Novice's Guide to ERIC - the Data Base of Education." Online 1, 3 (July, 1977): 24-29.

house interest. Whenever computer-based demonstrations and workshops are given or staff are being trained for online searching, we hope this manual will be useful. The original questions loaded into the ONTAP file did not include questions covering every clearinghouse's scope of interest. We at ERIC/IR have solicited search requests from five clearinghouses and will try to incorporate the questions and "perfect answer sets" into DIALOG's as possible.

As an aid to all ERIC online searchers, we have included in the appendix several search saves for population level and tool concepts. These can be accessed and exchanged between DIALOG workspaces by using the procedures explained in the introduction to the appendix. As more search saves are developed for online searching, we will announce them in the ERIC/IR Occasional Newsletter.

Comments about this manual are welcome and should be addressed to:

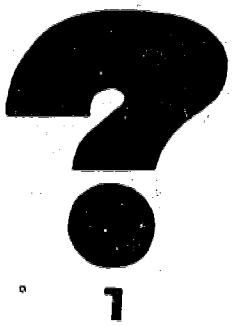
Pauline Atherton
ERIC/IR
Syracuse University
Syracuse, New York 13210

SECTION I

THE PROCESS OF ONLINE SEARCHING

INTRODUCTION TO THE STEPS AND FUNCTIONS OF ONLINE SEARCHING

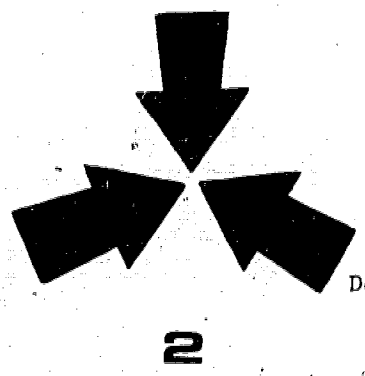
Online searching has been divided into eight steps, each requiring offline activities in preparation for the online activity. As the searcher progresses from one step to the next, he/she is faced with decisions which affect the succeeding steps. The following outline is accompanied by an abstract of the functions performed during each step in the process of online searching. The eight steps are treated in detail in the sections following the outline:



DEFINING THE INFORMATION-NEED AND SEARCH OBJECTIVE

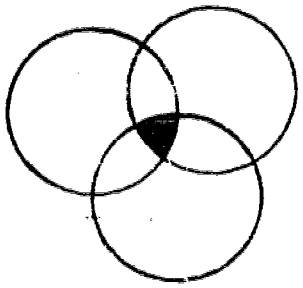
Interviewing the information seeker clarifies the request and determines search objectives:

a) retrieve all relevant items (high recall), or
b) retrieve only relevant items (high precision), or
c) retrieve some relevant items (Briefsearch).



IDENTIFYING RELEVANT DATA BASES

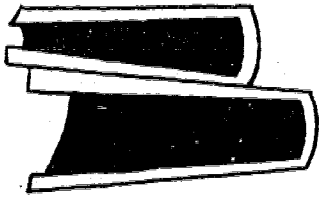
Determining which online data base to use first, which next, etc.



3

FORMULATING BASIC SEARCH LOGIC --
PLANNING SEARCH STRATEGIES

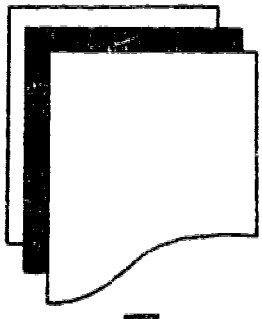
Analyzing the search topic into facets or concept groups. Planning approaches to search strategy for combining concepts of the topic.



4

COMPILING THE SEARCH TERMS

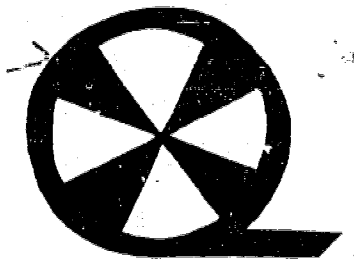
Choosing descriptors from the ERIC Thesaurus or other printed word lists. Selecting terms in identifier field and free text for searching subject-conveying fields (title, abstract, etc.). Using thesaurus and alphabetic word lists online.



5

ORDERING OUTPUT

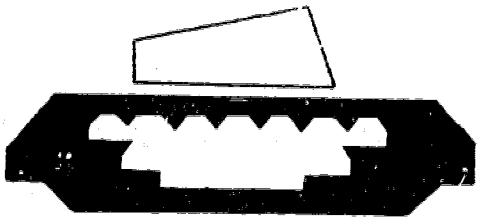
Limiting and printing output (offline and online). Selecting an approach to search strategy which best satisfies the search objectives expressed by the requestor.



6

CONCEPTUALIZING THE SEARCH AS INPUT TO THE RETRIEVAL SYSTEM

Arranging the search terms into concepts or facets, noting most important and less important concept groups, restricting or limiting output.



7

EVALUATING PRELIMINARY RESULTS

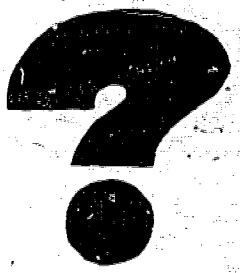
Reviewing search results, step by step, and considering alternate search strategies to meet search objectives (recycling steps 1-6).

8

EVALUATING FINAL RESULTS

Determining client satisfaction with search results.

STEP 1



DEFINING THE INFORMATION-NEED AND SEARCH OBJECTIVES

The first step in the search model is to identify the client's information-need through personal or telephone interviews or through written messages. Since this activity forms the basis of the search strategy, it is necessary that the information-need be described in as precise, specific and complete a manner as possible.

To ensure that the information specialist has received from the client all that is needed for searching, it may be advisable to check whether the following questions have been answered:

Clarification of Request	* Is there an agreement on the narrative statement of the information-need?
Request Negotiation	* Has the search objective been agreed upon? * Have the subject area terminology and literature sources been discussed? * Are there any limits or constraints to be imposed on output (date, language, type of information item)?
Search Vocabulary	* Do the client and information specialist agree upon the selection and/or exclusion of search terms? * Have the options of free text and controlled term searching been discussed? * What vocabulary aids have been selected?
Search Strategy Formulation	* Is there agreement on the formation of concept groups? * Have order of search statements and combinations of concept groups been discussed? * What alternate search strategies have been discussed? * Have output formats and limits been decided?
Administrative Details	* What are the arrangements concerning delivery of output and billing? * Has a post-hoc evaluation been arranged?

For any given search there appear to be at least three different types of search objectives:

1) to retrieve all relevant items -- a high recall search

2) to retrieve only relevant items -- a high precision search

3) to retrieve some relevant items -- a brief search

a) A high recall[1] search is formulated when the information seeker needs to find everything on the stated topic. It is necessary that the person writing a dissertation, or conducting a patent search finds all the relevant citations, thus a high recall search is required. In this case, the searcher must include all word variants and synonyms for each concept. An example of taking into account variant spellings and synonyms is taken from the search topic, parapsychology, the sample self-improvement exercise in Section II.

S PARAPSYCH?	Using the DIALOG truncation symbol, ?, allows the searcher to retrieve variant forms of the root, e.g., parapsychology, parapsychological, etc., instead of keying-in these terms separately.
S EXTRA(W)SENSORY S EXTRASENSORY	The term, extrasensory, a concept synonymous with parapsychological, is typed as a single word and as two adjacent words in order to retrieve different spellings of the same term.

Oftentimes a certain volume of non-relevant or marginally relevant output would have to be examined to ascertain that high recall had been obtained.

b) A high precision[2] search retrieves much relevant material, but with fewer non-relevant items in the output. Such a search would usually use only specific descriptors or terms representing concepts in the search with no generic term searching. An example of a high precision search formulation, the search topic "Identification of the Gifted" contains two concepts which can be represented by ERIC descriptors; thus, statements of a high precision search would include the

---

[1] Recall is defined as the percentage of relevant items in the file that are retrieved by the search.

[2] Precision is defined as the percentage of retrieved items that are relevant to the search topic.

descriptors, IDENTIFICATION and GIFTED. The following example would retrieve only ERIC accessions whose title includes both terms, GIFTED and IDENTIFICATION, or accessions with both IDENTIFICATION and GIFTED assigned as full descriptors.

```
S  GIFTED/DF,TI
S  IDENTIFICATION/DF,TI
C  1 AND 2
```

Thus, in the example above, high precision is dependent upon the function of the descriptors and the title to describe accurately the intellectual content of the document.

c) The brief search (called Briefsearch in this workbook) is done in response to the need for retrieving a few items either to lessen expenses or to perform a rapid survey of the file before a more comprehensive and lengthy strategy. The Briefsearch is necessarily a low recall search, performed in such a way that maximum results can be achieved with little online time. Choice of priority concepts is a factor in designing a Briefsearch.

A good example of the Briefsearch is in Search Topic #6, the Revision of the Anglo-American Cataloging Rules. In a single statement, using DIALOG's word proximity feature, the search is performed.

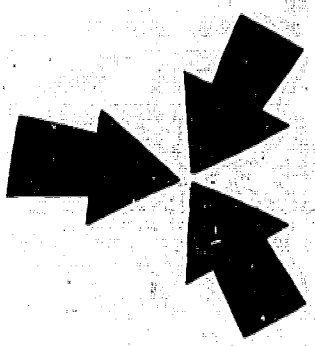
```
S  ANGLO(1W)CATALOGING(C)REVISION
```

The use of word proximity "(1W)" accounts for spelling variations and the use of the full-text searching capability (C) eliminates the necessity of entering additional SELECT and COMBINE operations.

More detailed formulations of the topic are given in Section II; as you proceed through the self-improvement exercises, compare the performances of the Briefsearch with the other formulations in terms of recall, precision, and online connect time. In most cases, the search queries in Section II contain sample formulations for each of these three types of search objectives.

STEP 2

IDENTIFYING RELEVANT DATA BASES



With over 150 bibliographic data bases online at this writing, this step in the search process is becoming more and more important. For the purposes of the manual, however, it is assumed that ERIC is the first choice and we will just move on.

For help with this step, the reader is referred to such articles as:

Donati, Robert. "Selective Survey of Online Access to Social Science Data Bases." Special Libraries 68, 11 (Nov. 1977): 396-406.

Donati, Robert. "Survey of Online Access to Social Science Data Bases." Paper presented at the Special Libraries Association Annual Conference, Denver, CO, June 8, 1976. 21p. (ED 128 011, $.83 microfiche, $1.67 photocopy, plus postage, available from EDRS).

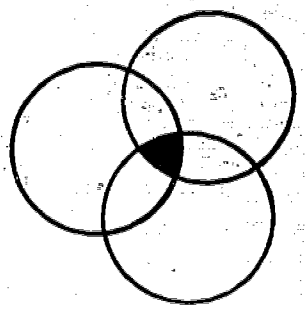
Hawkins, Donald T. "Multiple Database Searching: techniques and pitfalls." Online 2, 2 (April 1978): 9-15.

Teitelbaum, Henry H., and Donald T. Hawkins. "Database Subject Index." Online 2, 2 (April 1978): 16-21.

Wanger, Judith. "Multiple Database Use." Online 1, 4 (Oct. 1977): 35-41.

STEP 3

FORMULATING BASIC SEARCH LOGIC -- PLANNING SEARCH STRATEGIES



There is no single way to formulate a subject search, nor one best way given to fulfill the same search requirement. Different searchers will approach searches in different ways, and will assemble different formulations. Some of this is due to differences in the way people approach or attack a given problem in general. There may be several different formulations that give essentially the same performance, but all searchers will agree that a search should be analyzed into facets or concept groups. Facet indicates a separate aspect of a query. For example, a query on the effects of TV violence on children can be separated into three distinct concept groups (or facets, notions, parts, components):

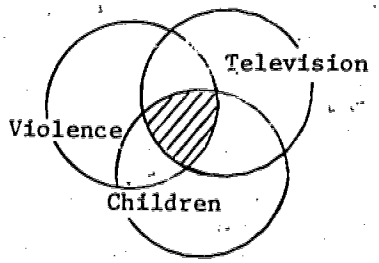
1) Television

2) Violence

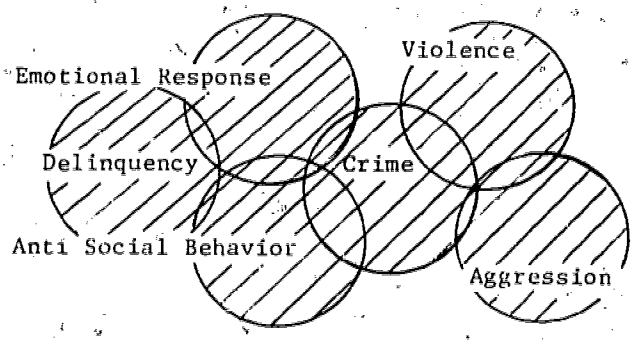
3) Children

In order to perform an online search, these three facets are joined using Boolean logic operators such as and, or, not. Since all three facets must be present to satisfy the stated information need, the Boolean operator "AND" is used with the search terms selected to represent the three facets.

In the illustration below, the three facets are represented by circles; the place where all three circles intersect forms the portion of the data base which contains records with all three facets.



In order to retrieve citations from the intersection of the three concept groups, each facet is represented by any number of synonymous terms. These terms are then assembled into a series of search statements and combined by the Boolean "OR" operator to produce a "complete" set of citations for that facet. In the example below, the facet VIOLENCE is made up of six search terms, crime, anti-social behavior, aggression, emotional response, violence, and delinquency, combined by the OR operator. How many of these terms and their variant forms could be searched depends on the searcher.



The inclusion of more synomous terms to the set representing VIOLENCE would relax the search requirement (or allow for more opportunity of matches) and permit the formulation to be satisfied by more citations. As a general principle which is independent of either data base or retrieval system, the inclusion of OR's increases the volume of output. On the other hand, the inclusion of AND's decreases the volume of output since every AND condition added to a formulation tends to make the search more restrictive (or reduce opportunities of matches) and results in less output. Thus both the total number of output citations, and the total number of relevant citations retrieved tend to decrease with each additional AND condition.

Likewise, the inclusion of NOT's decreases the volume of output. However, the NOT condition may also cause relevant citations to be excluded from the output when a citation contains the desired terminology and the unwanted terms. Thus, both the total number of output and the total number of relevant citations

decreases with each additional NOT condition.

It is quite possible that a given search request might be viewed by two different searchers as consisting of a different number of facets. This seems to depend upon the language of the topic, indexing vocabulary, and interpretation of the query and indexing language by the searcher when interviewing the client. For example, the search topic, "AV Aids for Library Instruction of Users," has been divided into two formulations in the following example:

Formulation A			Formulation B	
AV Aids	Library Instruction	Users	AV Aids	Instruction of Library Users
(1)	(2)	(3)	(1)	(2)

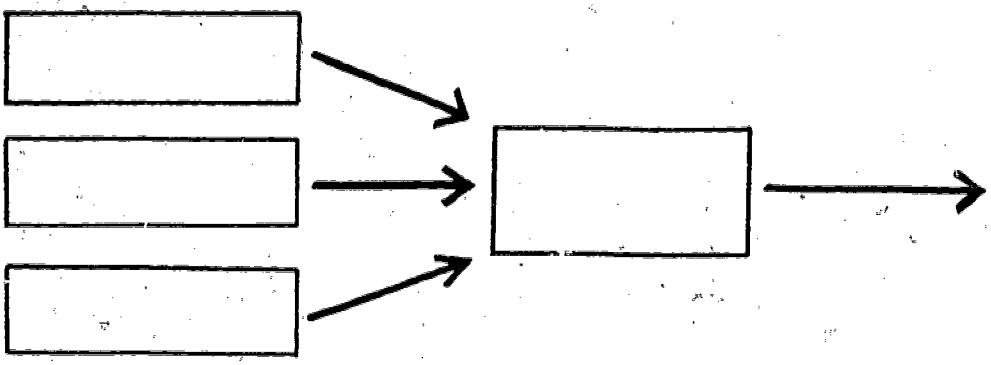
These differences in conceptualizing the query will result in different search formulations. The approaches to representing each concept in a query and the combination of conceptual groups form the central core of the online search strategy process. Models of several of the more predominant approaches to executing search strategy are shown in idealized representations in the following discussion on approaches to search strategy. No particular sequence is intended, and there is no intention to recommend one approach over the others. The object of the illustrations and discussion is to heighten each searcher's interest in which approaches are used and to aid in the analysis of which approach yields what kind of results.

1. Building Blocks

The first of the approaches is the building block (or sub-assembly, or modular) approach, and is characterized by an attempt to develop each aspect of the search separately, as if it were a subsearch all on its own, and then make the final logical assembly of all of these subprograms.

This is a predominant style for many searchers. Individuals who are accustomed to break all of their problems into subproblems may use the building block

approach as a natural extension of their own method of tackling a specific problem.



BUILDING BLOCK APPROACH

The building block approach, illustrated above, is demonstrated in the following search topic on "Identification of the Artistically Gifted." The topic is broken down into three concept groups: GIFTED, IDENTIFICATION, and ARTISTIC. In the table below are listed the three facets accompanied by search terms used to represent the facets in the search formulation.

ARTISTIC	GIFTED	IDENTIFICATION
Art	Gifted	Identify
Arts	Talent	Identified
Artistic	Talents	Identifying
Artistically	Talented	Identification
Esthetic -ally		Identifies
Esthetics		
Aesthetic -ally		
Aesthetics		

Terms from each facet are then assembled into search statements joined by the OR operation. The final step in the building block approach is to combine the results of the subsearches, in this case, by the AND operator.

Formulation	Comments
1 S ART? ?	Search statements 1-4 form the concept group ARTISTIC
2 S ARTISTIC?	
3 S ESTHETIC?	
4 S AESTHETIC?	

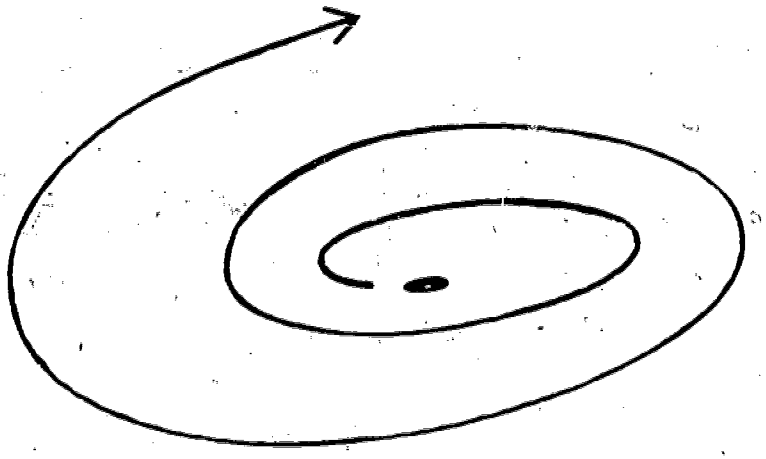
Formulation (cont.)	Comments
5 C 1-4/(	
6 S GIFTED	Statements 6 and 7 form the facet GIFTED
7 S TALENT?	
8 C 6-7/OR	
9 S IDENTIF?	Statement 9 forms the facet IDENTIFICATION
10 C 5 AND 8 AND 9	The combination of the three facets by the AND operator is done in the final search statement

The building block approach has the advantage of providing a somewhat clearer trace history, one that is easy to review and understand at a later date. It tends to follow and read like the actual query formulation; however, disadvantages of the approach are that more working memory and online time are used in comparison to the other approaches since the searcher establishes new sets with each command and introduces the main AND conditions only at the final stage of formulation. This approach may use more online time than is necessary. For example, the union of two concept groups in a 3-facet search may result in no or so few citations that it would not be necessary to impose the third condition. In that case, developing the third subprogram before a final COMBINE operation on the three subprograms would be an unnecessary use of time and effort. A modification of this approach would be to perform a Briefsearch in order to establish the likelihood that the combination of one term from each of the three facets results in some hits.

2. Citation Pearl Growing

This approach makes the most use of the interactive capabilities of the online search system. The technique here is to start with a very direct search on the most specific terms in the search request and find at least one citation. The searcher then calls up one or more of these citations for online review, noting index terms and free text found in some relevant citations. These new terms are

incorporated into the formulation to retrieve additional citations. After adding these terms to the formulation, one can again review more retrieved citations and continue this process in successive iterations until no additional terms are found that seem appropriate for inclusion.



CITATION PEARL GROWING APPROACH

The citation pearl growing approach to search strategy is illustrated above. The topic, "Identification of the Artistically Gifted," is searched using this approach. The Briefsearch is first employed to retrieve a few citations for online review.

Briefsearch

1 S GIFTED(C)IDENTIFICATION(C)ARTISTIC

One citation is retrieved and printed in format 2 so that the descriptors can be reviewed for possible inclusion in a revised formulation.

ED104039 EC071443
THE IDENTIFICATION OF ACADEMIC, CREATIVE AND LEADERSHIP TALENT FROM BIOGRAPHICAL DATA. FINAL REPORT.
INSTITUTE FOR BEHAVIORAL RESEARCH IN CREATIVITY, SALT LAKE CITY, UTAH.; NORTH CAROLINA STATE DEPT. OF PUBLIC INSTRUCTION, RALEIGH. DIV. FOR EXCEPTIONAL CHILDREN.
74 76P.
SPONSORING AGENCY: Z. SMITH REYNOLDS FOUNDATION, SAPELO ISLAND, GA.
EDRS PRICE MF-$0.76 HC-$4.43 PLUS POSTAGE
DESCRIPTORS: ACADEMIC ACHIEVEMENT/ *BEHAVIOR PATTERNS/ *CASE STUDIES (EDUCATION)/ CREATIVE ABILITY/ *CULTURAL FACTORS/ EXCEPTIONAL CHILD RESEARCH/ *GIFTED/ HIGH ACHIEVERS/ *IDENTIFICATION/ LEADERSHIP/ RACIAL FACTORS

Substituting the descriptor phrase, "Creative Ability," for the term artistic, the Briefsearch is revised.

Briefsearch

```
2 S CREATIVE(W)ABILITY(C)IDENTIFICATION(C)GIFTED
```

Five citations are retrieved; typing the retrieved citations in format 6 allows the searcher to browse the titles of the set in order to find candidates for printing more portions of the record.

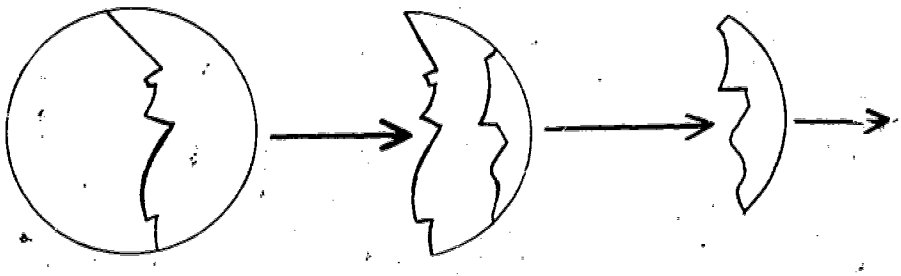
```
1/6/1
ED104095
  THE GIFTED AND TALENTED: A HANDBOOK FOR PARENTS. WORKING DRAFT.

1/6/2
ED104094
  THE IDENTIFICATION OF THE GIFTED AND TALENTED.

1/6/3
ED104039
  THE IDENTIFICATION OF ACADEMIC, CREATIVE AND LEADERSHIP TALENT FROM
BIOGRAPHICAL DATA. FINAL REPORT.

1/6/4
ED102773
  SUGGESTIONS FOR IDENTIFICATION OF GIFTED AND TALENTED STUDENTS.

1/6/5
ED100102
  TEACHING GIFTED CHILDREN ART IN GRADES ONE THROUGH THREE.
```

The last title seems relevant to the search topic so it is then printed in format 2.

```
ED100102  95  EC070972
  TEACHING GIFTED CHILDREN ART IN GRADES ONE THROUGH THREE.
  LUCA, MARK C.; ALLEN, BONNIE
  CALIFORNIA STATE DEPT. OF EDUCATION, SACRAMENTO. DIV. OF SPECIAL
EDUCATION.
  74   46P.; FOR ADDITIONAL INFORMATION, SEE ED 088 253 AND 254 AND ED
082 433
  SPONSORING AGENCY: BUREAU OF ELEMENTARY AND SECONDARY EDUCATION
(DHEW/OE), WASHINGTON, D.C.
  EDRS PRICE MF-$0.76 HC-$1.95 PLUS POSTAGE
  DESCRIPTORS: *ART/ CLASS ACTIVITIES/ CREATIVE ABILITY/ CREATIVE
EXPRESSION/ *CURRICULUM DEVELOPMENT/ CURRICULUM GUIDES/ EXCEPTIONAL
CHILD EDUCATION/ *GIFTED/ PRIMARY GRADES/ *PROGRAM PLANNING
  IDENTIFIERS: ELEMENTARY SECONDARY EDUCATION ACT TITLE V/ ESEA TITLE,
```

Additional searching vocabulary derived from this record may include "Creative expression," "art" or "arts." This process could then be continued by adding new terms to the Briefsearch formulation or incorporating search terms into a building block or other approach.

Citation pearl growing is characterized by the development of the search formulation in a very dynamic, empirical manner. Thinking time associated with this approach may result in a longer online connect time than might be achieved with searches in which some preparatory work and planning is done before online searching. Not only can online review of retrieved citations be very helpful in the identification of search terms for addition or deletion, but the payoff from this approach can be incorporated into any of the other approaches.



SUCCESSIVE FRACTIONS APPROACH

3. Successive Fractions

Successive fractions (or divide and conquer, or file partitioning) implies that an initial bite of the file is made in order to assemble a set that satisfies the conditions of the first facet of a multi-faceted search. This first facet may be a portion of the file such as year of publication, accession range, ERIC clearinghouse, document type or availability, major descriptor or major identifier. When the second facet of the search query is applied as an AND condition to the

partitioned subfile, the result is the establishment of an even smaller set or subfile. The remaining search facets continue to be applied to the subfile resulting from the previous AND conditions.

DIALOG's LIMITALL feature, when used at the beginning of a search, serves as an example of the successive fractions approach. In the following example, the LIMITALL capability is employed before the search formulation in order to restrict searching to the partitioned subfile, which is, in this case, only the periodical literature subfile.

Formulation	Comments
LIMITALL/EJ	
1 S IDENTIFICATION(F)GIFTED/TI	The retrieved set will only contain the terms "Identification" and "Gifted" in the title of a journal article.

The LIMITALL feature can also be used to restrict document year by using the accession number range. Since ERIC ONTAP contains only 1975 ERIC accessions, the LIMITALL capability for document year has already been imposed.

The successive fractions approach is also employed to restrict volume of output. A good example of this technique is found in search topic #3, Evaluation of Primary School Reading Materials. Since the combination of all three concept groups, READING, EVALUATION, and PRIMARY EDUCATION retrieves over 300 citations, the volume of output is restricted by limiting search terms to occurrences in the title, identifier, and descriptor fields. Output can also be limited by clearinghouse, document availability or type.

Using the successive fractions approach, the searcher can work with sets that become progressively smaller, and the search can be terminated whenever the size of the set is satisfactory to the searcher. For example, in the following search formulation, 45 citations are retrieved after the application of the second

facet. Instead of imposing the third facet, ARTISTIC, and risking the loss of some relevant items, the searcher can terminate the search at this point or limit the output, as in this case, by clearinghouse.

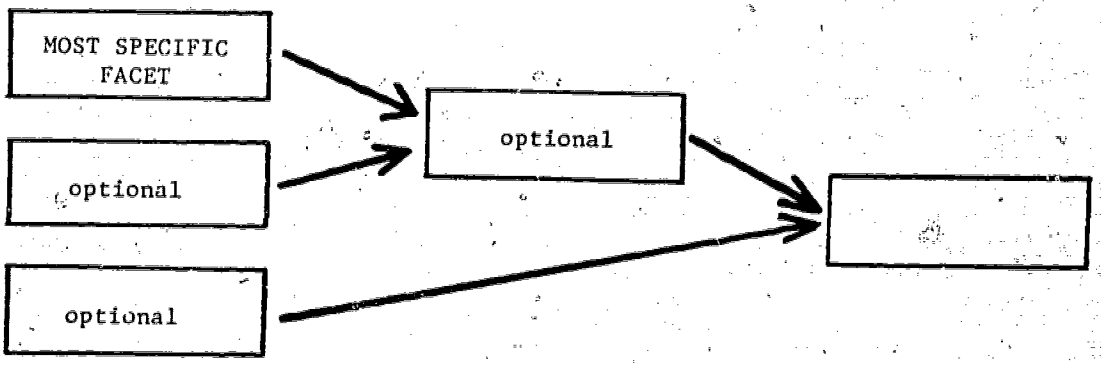
Formulation	Comments
1 S IDENTIF?	
2 S TALENT?? ?	
3 S GIFTED	
4 C 1 AND (2 OR 3)	45 citations are retrieved at this point
5 S CH=EC	Rather than impose the third facet to the set, the searcher restricts the output to Exceptional Children Clearinghouse. The output contains 26 citations at the conclusion
6 C 4 AND 5	

The major advantage with this approach is that it permits the search to be completed at an earlier point than would be the case with the building block approach if the size of the retrieved set fulfills the search objective before exhausting the search facets. The successive fractions approach also eliminates the necessity of "backtracking" or reformulating the search strategy when the final set fails to satisfy the search objectives.

The Boolean NOT operator is used in the formulation below to demonstrate how the successive fractions approach is employed to create a facet AMERICAN INDIANS; since the Indians of India are not wanted in the retrieved set, it is easier to use the NOT operator rather than keying-in separately all American Indian tribal names.

Formulation	Comments
1 S INDIAN	Statements 1 to 3 create a set containing all occurrences of the terms Indian or Indians. Truncation is not used because of the likelihood of retrieving false drops like Indiana or Indianapolis.
2 S INDIANS	
3 C 1 OR 2	
4 S INDIA	Statement 4 creates a set containing all occurrences of the term India. It is expected that documents containing this term would treat the subject of the Indians of India, a topic not desired.
5 C 3 NOT 4	The set of documents in statement 4 is "subtracted" or partitioned from the set created by statement 3.

The NOT operator serves as a good illustration of how the file is virtually partitioned and attention is paid only to the newly assembled portion in the successive fractions approach. This approach is often used advantageously with files such as BIOSIS, CA Condensates and Agricola as these files have major section headings that are used in the printed indexes, which when used as search terms at the beginning of a search, form a subfile for detailed subject searching.



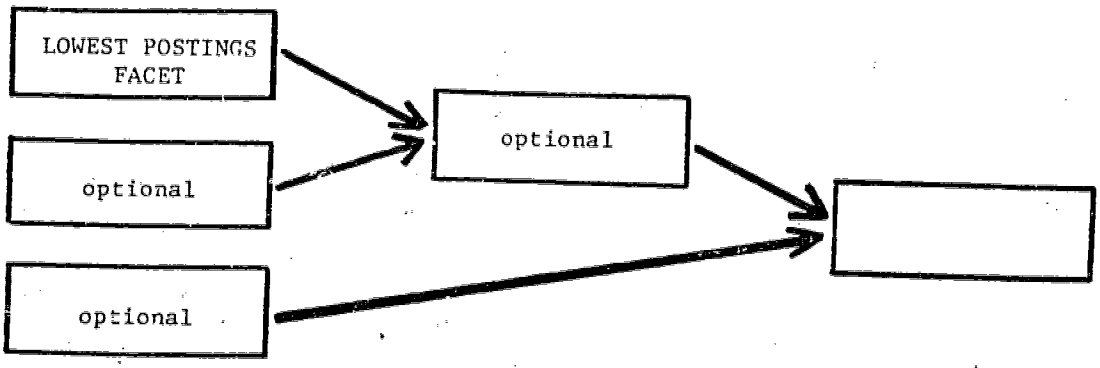
MOST SPECIFIC FACET FIRST APPROACH

4. Most Specific Facet First

This fourth approach consists of starting a multi-faceted search formulation with the most specific aspect of the query that is not likely to suffer from any vagueness of indexing (e.g., a named object such as Foreign Language Test, or Boy Scouts). If the postings are sufficiently low after scanning the results, the searcher may choose not to impose the remaining conditions to the search unless high precision performance is a search objective.

In search topic #2, "4-H Clubs, their members and activities," the search formulation is comprised of variant spellings of the club. Since 21 citations are retrieved using free text and 14 citations searching the identifier field, the search is terminated without applying the "members" or "activities" aspect of the search as the final set, in both cases, is sufficiently small.

If there had been too many citations after searching the most specific facet, then the next most specific facet would be searched and combined in an AND relationship with the results obtained with the first facet. The resulting intermediate result may now be small enough that there is no need to impose the rest of the search conditions. This approach can be more efficient than the full building block model when a search topic has some very specific facets with low postings.



LOWEST POSTINGS FACET FIRST APPROACH

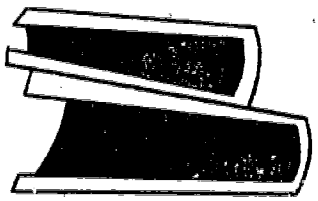
5. Lowest Postings Facet First

The fifth approach starts a multi-faceted search with the facet that is estimated (e.g., from the term frequency search aids, 7th edition ERIC Thesaurus, or from experience) to have the smallest number of postings. If the postings are sufficiently low there may be no reason to continue searching additional facets. For example, a searcher may safely assume that the facet GIRL SCOUTS OF AMERICA derived from a topic on "Success rates in Higher Education of students who were members of the Girl Scouts," has fewer postings that the facet HIGHER EDUCATION. Since the search for the facet GIRL SCOUTS results in less than 20 citations, the searcher terminates the search before applying the second facet HIGHER EDUCATION to the formulation.

Summary

There are other approaches to interactive searching, and of course, it is always possible to work with combinations of the above approaches. There is little in the way of solid theory or guidelines regarding how a search should be formulated.

STEP 4



COMPILING THE SEARCH TERMS

Before going online, some check of word lists and thesauri will help insure that the search terms used will lead to useful results. Both the controlled vocabulary of a data base (indexing records in descriptor/identifier fields) and the free text (subject-conveying fields other than descriptor and identifier fields) are possible sources of search terms.

1. Search Terms in Descriptor and Identifier Fields

In the ERIC data base, the descriptor and identifier fields contain terminology which has been assigned by ERIC Clearinghouse staff. The Thesaurus of ERIC Descriptors is an example of a controlled vocabulary which is used by the indexer in ERIC clearinghouses. If the Thesaurus is also used by the searcher the intellectual effort required to compile search terms can be reduced as it serves as a control on variant word forms, synonyms and homographs. For example, the ERIC Thesaurus suggests a number of search terms to represent the concept of violence: crime, aggression, antisocial behavior, delinquency, emotional response, and the descriptor, violence.

The identifier field contains names of persons, places, things, which are important terms associated with the item. No controlled list like the Thesaurus is used by the indexer when identifiers are noted, but there are some rules for entry which control this operation, and each type of identifier is now being reviewed and checklists compiled by the ERIC Network.

2. Free Text Searching

The free text searching capability of DIALOG allows for the scanning of all subject-conveying fields of the ERIC record: title, abstract, descriptor, identifier, descriptive note, corporate source and sponsoring agency. If a concept like

parapsychology is not found in the ERIC Thesaurus, terms synonymous with parapsychology can be gathered from such lexicographic sources as the Thesaurus of Psychological Terms, Roget's Thesaurus, an unabridged dictionary and other sources. This list of search terms compiled from sources other than the ERIC Thesaurus can be searched as free text in all, or a specified number of subject-conveying fields; however, recall scores may be low if only a single subject-conveying field is searched.

As citations are retrieved using this free text scan method, a few citations can be called up for online review to note how relevant citations are indexed. These descriptors and other free text terms can be used in further search formulations. This method of gathering additional search terms resembles the "citation pearl growing" approach to the formulation of search logic discussed in the previous step of this section. Titles of known relevant items are also a rich source of "free text" search terms.

3. Online Alphabetic Word List and Thesaurus

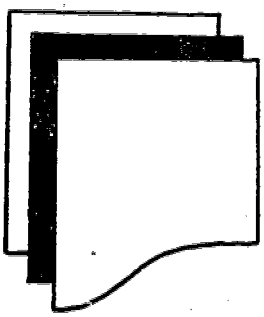
When a planned search has more than one term associated with it that occurs close together alphabetically, it is fast and convenient to review the online dictionary (an alphabetic list of words in all the searchable fields of the ERIC data base). Rather than directly keying-in the search terms in a series of SELECT commands the EXPAND command is used. Additional terms may appear which had not been considered previously.

Entries like corporate sources and author names are particularly appropriate author and corporate source names and because typing errors may occur when keying-in long corporate source names.
corporate source names.

Scanning the online thesaurus, the subject-related descriptor list, may suggest to the searcher during the online activity search terms which do not appear in alphabetic proximity. The postings for each term are given and this

data can affect the search strategy being considered. Heavily posted terms may be selected after reviewing the EXPAND display of a lightly posted term, for instance, or vice versa. These strategies are discussed in the previous step. However, selecting a phrase from the online thesaurus or alphabetic word list restricts retrieval to occurrences of the phrase as a multi-term descriptor or identifier. Retrieving single and multiple word search terms from all subject-conveying fields is covered in more detail in step 6.

31



STEP 5

ORDERING OUTPUT

Ordering output occurs not as a single, isolated event in the sequence of online search activities. Rather, because decisions influenced by output specifications are made throughout the search process, the actual activity of ordering output is not performed at any fixed moment during the online search. As soon as the searcher begins formulating search strategy, the construction is affected by conditions imposed by the requestor concerning output specifications, e.g., size of output, document type, year of publication, etc.

For example, in order to fulfill the request of an information seeker who desires a small amount of output with a high percentage of relevant items, the searcher may select the "citation pearl growing" approach to search strategy. The LIMITALL command may be preferred by a searcher who must restrict the results by publication date. In another instance, the Briefsearch and a shortened online print format may be favored when the information seeker is hampered by time restrictions and requires immediate results; a more comprehensive search may be performed at a later time.

Just as ordering output influences the planning of search strategy (step 3 in the process of online searching), it affects the conceptualization of the search as input to the retrieval system (step 4). If the information seeker is concerned about retrieving <u>all</u> relevant citations at a minimum cost, the searcher is compelled to select an approach which will result in high recall, and formulate the strategy employing time-saving devices, e.g. EXPAND and range SELECT, low activity search time, pre-search offline preparation to minimize online "think" time, etc.

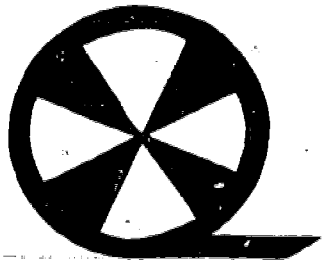
Output can be scanned online for fine tuning or for further trimming of the search results on the basis of non-subject parameters. For example, the resultant set may exceed the output size requested by the information seeker so that the searcher, depending upon policy, personal judgment or consultation with the requestor may decrease the size by applying limiting criteria such as language, date or type of publication, availability, sponsoring agency, etc. These specifications are in most cases applied to an existing set. The decision to trim the output may be the outcome of the preliminary evaluation (step 7), in which the search results are reviewed to determine whether the search strategy satisfies the search objective.

The retrieval system and/or data base plays a role in determining characteristics of the output. Sorting criteria vary among data bases; format of output not only differs from DIALOG data base to data base, but DIALOG offers predetermined output formats rather than allowing the searcher to structure his/her own format.

In summary, search output can be limited by criteria imposed by the information seeker, searcher, or retrieval system. The influence of ordering output prevails throughout the search process; however, ordering output, as a step in the process of online searching, exists as a concrete and very real activity, but may occur at a number of points during online interaction.

STEP 6

CONCEPTUALIZING THE SEARCH AS INPUT TO THE RETRIEVAL SYSTEM



The features of any computerized retrieval system influence the way in which a given search is formulated for that system. Only a few of the operations which can seriously affect search strategy operations are described briefly below. For a more complete description, see A Brief Guide to DIALOG Searching, and C. Bourne, DIALOG Lab Workbook.

1. Word Proximity Searching for Single Term Combinations

The word proximity feature can be used as a shortcut to express a direct logical AND combination. For example, a Briefsearch employing word proximity would achieve the same search results as the more conventional, sequential, formulations in the following search topic on the effects of TV violence on children:

Sequential Search Formulation

1 S CHILDREN
2 S TELEVISION
3 S VIOLENCE
4 C 1-3/AND

Word Proximity Search Formulation

1 S CHILDREN(C)TELEVISION(C)VIOLENCE

There are some disadvantages with the word proximity approach. It does not provide any intermediate sets. In the event that there is an interest in modifying the search formulation, it cannot be used with truncated words.*

The use of the word proximity feature with its implicit AND operator is the preferred practice for searches, or parts of searches, when:

---

* This will be changed in DIALOG very soon.

1) the resulting command statement is a single typed line in total length, and

2) the search terms are lightly posted.

Word proximity is particularly advantageous when the search is planned to be made over all rather than a few specified fields.

2. Multiple-Word Search Terms -- Word Proximity Searching within a Field of the Record

In the case of single word search terms, DIALOG scans all subject-conveying fields by default unless otherwise directed by the searcher. Multiple-word terms entered by the searcher are only checked against the entries from the descriptor or identifier fields. Searchers restricting their searches to the controlled vocabulary fields of the ERIC record (descriptors and identifiers) are ignoring other subject-conveying items of the record. High recall is achieved when the searcher uses both the controlled vocabulary and free text. Searching can be extended to all subject-conveying fields simply by using the adjacency feature (W) as shown below.

Search Formulation	Result	Effect
S BOOK CATALOGS	85	Searches in descriptor and identifier fields for subject headings that include the character sequence, BOOK CATALOGS
S BOOK(W)CATALOGS	95	Searches as above, plus title, abstract, descriptive note, corporate source, and sponsoring agency fields for the occurrence of this character string

The use of the word proximity feature in this way generally retrieves about 20% more postings than are retrieved by searching fields with assigned index terms. The significant increase in postings results in a high recall search. But remember, more computer time is spent using word proximity, especially for high posting combinations like ELEMENTARY(W)EDUCATION.

There are many instances in which the controlled vocabulary does not yet include current expressions that may appear in the literature. The word proximity feature can help find these citations by means of direct text searching as in the following examples:

A. Search by Controlled Terms	Results A	Results B	B. Word Proximity Free Text Search
S TAXPAYER REVOLT	0	4	S TAXPAYER(W)REVOLT
S INFORMATION EXPLOSION	3	40	S INFORMATION(W)EXPLOSION
S STREET PEOPLE	1	4	S STREET(W)PEOPLE
S BACK TO BASICS	12	59	S BACK(1W)BASICS
S GENERATION OF ALTERNATIVES	0	1	S GENERATION(1W)ALTERNATIVES

3. Storage-Saving Techniques to Bypass Computer Memory Limitations

The search formulation can be influenced by the amount of working storage available to the searcher. About one million postings per search are available to each searcher. Computer memory limitations pose less of a problem in the ERIC file than in large files such as BIOSIS or AGRICOLA. Since problems do occur, certain formulation strategies can be used to reduce the amount of storage space needed. Storage-saving techniques that can be used are:

a) Avoidance of intermediate logical formulations

b) Stem search

c) EXPAND and range SELECT

d) Word proximity feature

e) LIMIT feature

f) BEGIN, resetting set numbers to zero

g) Search Save

a) Intermediate results of COMBINE operations that are stored online take up space in the same way as the other sets. Combining all the logic in one statement reduces storage space. Consider the following example in which

the same search objective is met two different ways, but with differences in the total amount of working storage space required:

Search Plan: (TEST? or EVALUATE? or ANALY?) and (READING or MATHEMATICS)

Formulation A

```
1   38050   S TEST?
2   49153   S EVALUAT?
3   47022   S ANALY?
4  100632   C 1-3/OR
5   21461   S READING
6   10520   S MATHEMATICS
7   30507   C 5-6/OR
8   14598   C 4 AND 7
```

(311,943 postings in working storage)

Formulation B

```
1   38050   S TEST?
2   49153   S EVALUAT?
3   47022   S ANALY?
4   21461   S READING
5   10520   S MATHEMATICS
6   14598   C (10R20R3)  AND  (40R5)
```

(180,804 postings in working storage)

Performing the same search two different ways requires a difference of 131,139 postings in working memory, which could be important for some searches. The preferred practice is to avoid the intermediate logical operations if a storage overflow problem is anticipated.

b) If a word is sufficiently unique, the use of a word stem instead of the combination of a series of terms might save a small amount of storage space, because an intermediate logical operation might be avoided through the implicit OR operation performed in the stem search. Consider the following example in which the same search objective is met two different ways:

Formulation A

```
1   3874   S EVALUATE
2   3680   S EVALUATED
3   4024   S EVALUATING
4  42630   S EVALUATION
5   2373   S EVALUATIONS
6   2016   S EVALUATIVE
7  48737   C 1-6/OR
```

(107,334 postings in working storage)

Formulation B

```
1   49153   S EVALUAT?
```

(49,153 postings in working storage)

In formulation B, over 58,000 storage spaces have been saved by using stem searching. The preferred practice is to use the word stem when the word is sufficiently unique not to cause too many false drop problems.

c) The EXPAND and range SELECT feature can be very helpful with storage problems because of the implicit OR operation that is done with the range SELECT feature. This can be done for both the alphabetically-related terms and the thesaurus-related terms. In the following example, the same search objective is accomplished in two different ways:

Search Plan: (AUDIOVISUAL EQUIPMENT, FACILITIES) and (LIBRARIES of all types)

Formulation A

1	6,889	S	TELEVISION
2	4,796	S	FILMS
3	4,343	S	AUDIOVISUAL AIDS
4	2,694	S	INSTRUCTIONAL MEDIA
5	1,556	S	INSTRUCTIONAL AIDS
6	1,102	S	AUDIOVISUAL INSTRUCTION
7	996	S	SLIDES
8	995	S	TAPE RECORDINGS
9	800	S	INSTRUCTIONAL FILMS
10	469	S	EXHIBITS
11	426	S	VISUAL AIDS
12	338	S	PHONOTAPE RECORDINGS
13	265	S	PROJECTION EQUIPMENT
14	249	S	CARTOONS
15	226	S	PHONOGRAPH RECORDS
16	191	S	TAPE RECORDINGS
17	152	S	VIDEO CASSETTE SYSTEMS
18	79	S	SCREENS
19	75	S	REALIA
20	52	S	DISPLAY PANELS
21	40	S	MICROPHONES
22	19,040	C	1-21/OR
23	6,802	S	LIBRARIES
24	2,310	S	LIBRARY SERVICES
25	1,129	S	PUBLIC LIBRARIES
26	1,173	S	UNIVERSITY LIBRARIES
27	1,045	S	INSTRUCTIONAL MATERIALS CENTERS
28	801	S	SCHOOL LIBRARIES
29	740	S	INFORMATION CENTERS
30	529	S	SPECIAL LIBRARIES
31	371	S	ARCHIVES
32	289	S	STATE LIBRARIES
33	230	S	RESEARCH LIBRARIES
34	197	S	NATIONAL LIBRARIES
35	7,985	C	23-34/OR
36	1,048	C	22 AND 35

(69,478 postings in working storage)

Formulation B

		E(AUDIOVISUAL AIDS)
1	20,016	SR1-R18,R20-R31,R35-R37,R41
		E(LIBRARIES)
2	9,241	SR1-R13,R15,R22-R24
3	1,469	C 1 AND 2

(30,726 postings in working storage)

In this example the EXPAND/SELECT approach saved almost 40,000 postings in working memory. Moreover, formulation A took more time to type, increasing online connect time and increasing the possibility of typing errors.

d) The word proximity feature, discussed earlier in step 6, does not provide any intermediary sets. If the searcher intends to revise the formulation during the online interaction, using word proximity as a means to bypass computer memory limitations is not suggested. However, word proximity may be suitable to express a Briefsearch when there is no need for intermediate sets. Note in the following example on the effects of TV violence on children that over 47,000 postings are saved using word proximity.

Formulation A

```
1  38,189  S CHILDREN
2   7,930  S TELEVISION
3     984  S VIOLENCE
4     133  C 1-3/AND
(47,236 postings in working memory)
```

Formulation B

```
1  133  S CHILDREN(C)TELEVISION(C)VIOLENCE
(133 postings in working memory)
```

e) Use the BEGIN command when doing several unrelated searches at the same terminal session. This resets set numbers to zero, and clears all of your working storage space, thus starting subsequent searches with a full allocation of working storage.

f) The LIMIT feature can be used to do some tasks that might otherwise require more storage when done with set combinations. This is particularly true when restricting the search output to accession year or document type, as shown in the following example:

Search Plan: (VOCATIONAL TRAINING) and (ERIC accession YEAR = 1975 or more recent)

Formulation A

```
1  11,615  S  VOCATIONAL(W)EDUCATION
2  32,120  S  YR=75
3  37,145  S  YR=76
4  36,369  S  YR=77
5  12,317  S  YR=78
6   4,706  C  1 AND (2 OR 3 OR 4 OR 5)
(141,272 postings in working storage)
```

Formulation B: LIMIT Command

```
1  11,615  S  VOCATIONAL(W)EDUCATION
2   3,768     LIMIT1/095254-999999/ED
3   1,015     LIMIT1/101873-999999/EJ
4   4,783  C  2-3/OR
(21,181 postings in working storage)
```

In this case, the most convenient and fastest technique for the searcher is to use the YR= feature instead of looking up and keying-in the ED- and EJ- number ranges for reports. In the above case over 120,000 storage positions have been saved.

When storage problems are anticipated, the LIMITALL feature can be used to partition the file by accession number or document availability. The search is performed on the portion of the file specified in the LIMITALL command. Using the same search plan above, the LIMITALL capability is demonstrated below:

Search Plan: (VOCATIONAL EDUCATION) and (ERIC accession YEAR = 1975 or more recent)

Formulation C: LIMITALL Commands	Comments
LIMITALL/095254-999999/ED	The LIMITALL command restricts the file to 1975+ ERIC non-journal accessions.
1 3768 S VOCATIONAL(W)EDUCATION	Occurrences of the descriptor phrase are retrieved from this partitioned file.
LIMITALL/101873-999999/EJ	The new LIMITALL command cancels the first command.
2 1015 S VOCATIONAL(W)EDUCATION	
3 4706 C 1-2/OR	
LIMITALL/ALL	The LIMITALL command is removed in this way to permit searching in the entire file.
(9412 postings in working memory)	

Although the LIMITALL feature keeps the number of postings in the working memory at a minimum when comparing the results of the three strategies above, one disadvantage of the LIMITALL feature is that the searcher is required to repeat the search statements (in the above example, one phrase, VOCATIONAL(W)EDUCATION, is entered twice) if both the journal and non-journal citations are desired; however, LIMITALL is a powerful storage-saving device if one of the output specifications is year of publication or ERIC document type.

g) The search save is the most powerful storage-saving technique of all. A search save can be prepared and saved for frequently occurring facets and executed when needed. An example of its use is shown in the sample formulation below:

Search Plan: (READING or READER) and (ELEMENTARY or SECONDARY EDUCATION)

Search Save Formulation			Comments
1	22793	S ELEMENTARY(1W)EDUCATION	
2	30804	S SECONDARY(W)EDUCATION	
3	25949	S READING?	
4	4458	S READER?	
5	6898	C (3 OR 4) AND (1 OR 2)	
		END/SAVE	The Search Save command (END/SAVE) concludes the routine. DIALOG assigns a serial number to the Search Save and the searcher can resume searching.

SERIAL#1PAA
23MAY78 8:52:56 USER 4111
$0.00 0.046 HRS FILE1 00 DESCRIPTORS
$0.37 TYMNET
$0.37 ESTIMATED TOTAL COST

Search Plan: (READING or READER) and (ELEMENTARY or SECONDARY EDUCATION) and COSTS

Formulation			Comments
.EXECUTE 1PAA			The Search Save is recalled and performed by the system at the signal of the .EXECUTE command
	22793	ELEMENTARY(1W)EDUCATION	
	30804	SECONDARY(W)EDUCATION	
	25949	READING?	
	4458	READER?	
	6898	(3 OR 4) AND (1 OR 2)	The system responds to the .EXECUTE command by printing the results of all the search statements but only the final statement retains the record of postings. In this case, it becomes statement 1 of the search in progress.
1	6898	SERIAL# 1PAA	
2	6577	S COSTS	The third facet is applied to the formulation
3	91	C 1 AND 2	The search results

In the example above, the 90,902 storage locations used in statements 1-6 of the first formulation are erased and the searcher uses only 6898 storage locations for the two facets READING and EDUCATION.

Such a search tactic can also be temporarily constructed and used during a single terminal session. The Search Save has the effect of eliminating all temporary storage space for the development of the two facets READING and EDUCATION, and can then be released as soon as the rest of the search has been completed.

It should be noted that the Search Save approach can help with the storage problem, but extra computer time is required to reprocess all of the search steps

which comprise the saved search.

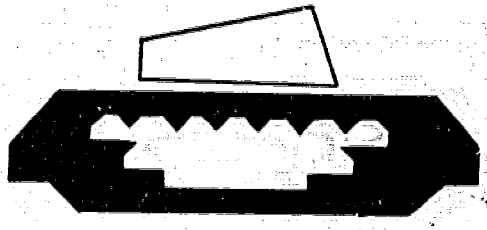
The appendix to this report includes some Search Saves for common educational level facets. The procedure for recalling and exchanging these Search Saves between workspaces is covered in the introduction to the appendix.

4. Searching during Low Activity Times

To reduce the cost of searching, the search should be performed during the system's low activity hours. Most online search systems have a response time, and hence a total search time, that depends upon the volume of processing activity that is taking place at that time. The greater the number of simultaneous searches, the slower the response time. Searches done during the 8-2 (PST) time slot are generally slower, since all U.S. time zones are in the course of their normal working day. Search speeds for the same search can sometimes differ by a factor of 2, depending upon the time of day that the search is made. A search that takes five minutes to run at 3 P.M. (PST), during the low activity period, could sometimes take 10 minutes to run at Noon (PST), a high activity period.

STEP 7

EVALUATING PRELIMINARY RESULTS



Fundamental to the evaluation of the search strategy is the question, "How well is the search formulation meeting the information need?" If it is determined in the original interview that the information seeker wants high recall, then an appropriate performance measure would be a check on the fraction of relevant citations in the file that are actually retrieved by the search. If the objective is only a few relevant items at a low cost to the information seeker, then search time would be an important measure and a Briefsearch may be satisfactory if some relevant items are retrieved. While online, the searcher should monitor the search in progress and check on results. This check may require a change in strategy and overall approach to the search. A review of Steps 1-6 may be necessary.

STEP 8

EVALUATING FINAL RESULTS

There is no single measure of how good a search is. Performance measures such as cost, recall, and precision can be applied to the search results given the specific search objectives:

Search Objective	Useful Performance Measures
* To retrieve all (or as many as possible) relevant citations in file	Recall (defined as percent of relevant citations in the file that are actually retrieved)
* To reduce (or keep low) the number of non-relevant citations retrieved	Precision (defined as percent of retrieved citations that are relevant)
* To keep costs low and to retrieve some relevant citations	Cost (defined as connect time, and cost of offline printing)

1. Recall

Recall is easy to define, but very difficult to measure because it is so hard to find out the total number of relevant citations that are in the file.*

Certain tasks can be performed in order to achieve high recall. A search done on both assigned index fields and titles will generally return more citations (total and relevant) than a search done on one of those fields themselves. The larger the total number of subject fields of the bibliographic records being searched, the larger the total number of returned citations.

2. Precision

Precision, sometimes called relevance, is easy to understand and easy to measure. Whereas recall increases as more parts of a citation are searched,

---

* The search questions provided in the next section have all been searched several times by several different people, and the output reviewed, so that there is a fairly good idea of what citations in the file are relevant to each of the questions. This information has been stored away as answer sets, so that it can be retrieved to permit the practice searcher to compute recall.

conversely, precision decreases as more parts of a citation are searched. If a low volume, high precision search is desired, searching the title, major descriptors and major identifiers delivers a high percentage of relevant items.

3. Cost

Connect time or computer cost is also easy to understand, and easy to measure because the accounting information is provided as part of each search output by the use of the BEGIN, END, .COST or LOGOFF commands. Using the time-saving schemes which have been described in Step 6 can serve to aid the searcher who desires to lower costs of online searching.

SECTION II

ERIC ONTAP EXERCISES FOR SELF-IMPROVEMENT OF ONLINE SEARCHING

INTRODUCTION TO DIALOG'S ERIC ONTAP (FILE 201)

ERIC ONTAP is a collection of RIE and CIJE citations from the 1975 ERIC file. Its contents correspond to citations in all 1975 printed ERIC indexes, not necessarily to a 1975 publication date, and include the range of ERIC accession numbers EJ 101 873 to EJ 121 926, and ED 095 254 to ED 110 594. ERIC ONTAP is exactly the same in all respects (data elements, searchable fields, etc.) as the 1975 accessions in the regular DIALOG ERIC file (file 1).

ERIC ONTAP represents 12% of the total ERIC data base (some 35,400 items). There are points to remember when using the file:

1. Since the file is small, you will not encounter problems involving size, e.g., what to do when you still have a large number of postings in your final set.

2. Seven editions of the ERIC Thesaurus were used in order to index file 1, but only the fifth edition was used for ERIC ONTAP indexing. Thus, shifts in meaning and usage of terms over the period of several years do not occur in ERIC ONTAP to the same extent as in file 1.

The unusual feature of the ERIC ONTAP file is that it contains "answer sets" (lists of ERIC accession numbers) to predetermined queries. Searchers using this file can perform a search for the query and compare their results with the "answer set" for that query.

The search queries represent real, not synthetic questions, since they are assembled from questions posed to reference librarians and ERIC Clearinghouse searchers. Level of difficulty varies from query to query and they are graded as simple, moderate, and difficult. "Named object" queries, e.g., 4-H Clubs,

16 Personality Factor Test, are easiest to perform, but they involve problems for the online searcher which, if worked out, will help in more complex search formulations.

This section of this manual includes 10 of the 29 queries in the ERIC ONTAP file. We have chosen queries of special interest to the 16 ERIC Clearinghouses. It is suggested that Clearinghouse searchers review some or all of SECTION II search topics rather than the single topic of interest to their clearinghouse in order to execute online search operations described in SECTION I. For demonstrations or workshops, searchers may want to use specific queries to show how online searching of the ERIC data base can be performed and to point out how differences in search strategy retrieve different output.

Each query in this section will include various search formulations with explanations. The way in which the searches are executed will vary somewhat from what would likely be seen in actual practice. This is done solely for clarity of presentation. It is assumed that the searcher will be working in an interactive manner, using EXPANDS, reviewing sample citations, and restructuring the strategy as working results are obtained. Sample search formulations are given here to show the critical elements of the final formulation and to emphasize how a change in search objective (i.e., high recall, high precision, Brief-search), can effect retrieval results.

In many cases, the search formulations given for each query are composites that were assembled with hindsight after viewing the results of several formulations. (Executing the "perfect search" at one sitting is still elusive.)

The "answer set" is a list of ERIC documents that are judged to be relevant to the topic, which are cited by ERIC accession number. The set can be selected online by using the answer set number, e.g., S AN=D05. Relevance judgments for each query are made on the basis of the total record in ERIC ONTAP, not on the basis of examination or analysis of the full document or article, and are open to question and modification. Broad review publications, e.g., bibliographies of

dissertations in education, are not judged as relevant unless there is specific mention of the search topic in the abstract.

Since the list of relevant citations, the "answer set," is compiled from the results of several searches on each topic, it represents an approximation to the total number of relevant citations contained in the entire ERIC ONTAP file. (There may be additional relevant citations in the file for most of these questions. We would appreciate hearing about them from those of you who find them.) The "answer sets" to search topics 11-15 are not contained in the ERIC ONTAP file but are given at the conclusion of each discussion.

For each query in this section, one search formulation, specifically designed as a high recall search, retrieves all the relevant citations listed at the conclusion of the discussion. In addition, one search formulation, in most cases, is singled out as being preferred in the light of the information need and search objective stated in the query description.

You, the reader/searcher, should try your own search strategies and compare your output and your choice of search vocabulary with that of the high recall search.

The Briefsearch is also given for almost every query along with an example of a high precision formulation.

All the approaches to search formulation, e.g., building block, most specific aspect first, etc., are covered in one or more queries. For example, the stated search objective accompanying the query on parapsychology requires high recall. Consequently, the formulation which achieves high recall is selected as the preferred routine. Do remember, however, that no single search strategy or approach to search formulation is deemed the best in all cases.

To obtain an explanation of the features of the ERIC ONTAP file, execute a BEGIN 201 command. Typing ?ONTAP produces the following directions:

```
ONTAP (ONLINE TRAINING & PRACTICE) FILE
PROVIDES TEST QUESTIONS & RELEVANT
CITATIONS FOR EACH QUESTION.  THE 1975
ERIC FILE IS THE SOURCE OF THE ANSWER
SETS.  QUESTIONS AND ANSWER SETS ARE
LABELLED AS TO COMPLEXITY, E.G. AN=S04,
AN=M04, AN=D03 DENOTE THE ANSWER SETS
FOR THE 3RD SIMPLE, 4TH MEDIUM, AND 3RD
DIFFICULT QUESTION.  AFTER FINISHING THE
SEARCH, THE PERFECT ANSWER SET FOR THAT
QUESTION CAN BE SELECTED BY COMMANDS OF
THE FORM 'SELECT AN=S03', AND SELECTED
ANSWER SETS CAN BE COMBINED IN AN 'AND'
RELATIONSHIP WITH YOUR FINAL SEARCH SET
TO FIND OUT HOW MANY RELEVANT CITATIONS
YOU FOUND.  A COMMAND /EVAL COMPUTES THE
RECALL AND PRECISION OF YOUR SEARCH.
SEE    ?TAPSAM FOR SAMPLE SIMPLE SEARCH
       ?TAPSIM FOR LIST OF SIMPLE QUESTS.
       ?TAPMED FOR LIST OF MEDIUM QUESTS.
       ?TAPDIF FOR LIST OF DIFFICULT
        QUESTIONS
       ?EVAL FOR EXPLANATION OF PREC/REL.
```

There are additional instructions which explain features of ONTAP in greater detail. Instead of listing these texts below, they are given as needed in discussions of the search queries which follow.

It is suggested that you start by trying the first query on parapsychology. Read the text, develop your own strategy, and perform an online search. Then figure the recall and precision ratios of your performance; this can be done by comparing your output with the list of ERIC accession numbers given in this workbook, or, selecting the answer set for the particular query and using the /EVAL command available in the ERIC ONTAP file. The /EVAL command will automatically compute your recall and precision ratios.

The discussions of the parapsychology query and the fifteen other topics are arranged according to steps 3 to 8 of the online searching process outlined in Section I. The steps are summarized below:

Step 3. Formulating basic search logic -- planning search strategies

Step 4. Compiling the search terms

Step 5. Ordering output

Step 6. Conceptualizing the search as input to the retrieval system

*Step 7. Evaluating preliminary results

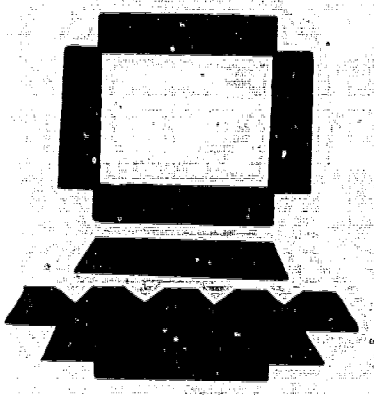
Step 8. Evaluating final results

*Since step 7 affects search formulation and occurs during the online interaction, the discussion of this step is treated, for the most part, in the Comments section of search formulations.

We recommend you begin the ONTAP exercises with the first query on parapsychology. The discussion of this query focuses on both the steps of the online searching process and the procedures for using ERIC ONTAP. The discussion of parapsychology will familiarize you with the features of the ONTAP file so that you can prepare your own search strategy. The remaining queries follow the same pattern of discussion, but the instructions for online features of the ONTAP file are not covered in as much detail in subsequent discussions.

SAMPLE SEARCH TOPIC

PARAPSYCHOLOGY



Topic: Parapsychology. (S AN=S09)

Search Objective: All information in ONTAP on this topic

Step 3. Formulating Basic Search Logic -- Planning Search Strategies

The search topic, parapsychology, is depicted by a single concept or facet. The term parapsychology, its variant forms and synonymous terms, e.g., ESP, psychic, extrasensory, are incorporated into the search strategy. All terminology chosen to represent the facet are joined together by the use of the OR operator, thus forming a single set to describe the concept.

Step 4. Compiling the Search Terms

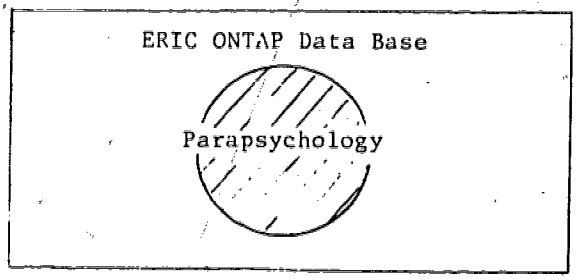
Usually the search terms will come from the ERIC Thesaurus; however, in this case, parapsychology is not used as an ERIC descriptor nor are related terms such as ESP, psychic, extrasensory, etc. Consequently, other aids for collecting searching vocabulary may be consulted such as Roget's Thesaurus or Thesaurus of Psychological Terms. Since no ERIC descriptors describe the concept of parapsychology, searching only the descriptor field of the ERIC record would be useless. The identifier field of ERIC resumés may contain these terms, as it is comprised of terminology which has been assigned to the document by ERIC clearinghouse staff in order to indicate the document's intellectual content.

Step 5. Ordering Output.

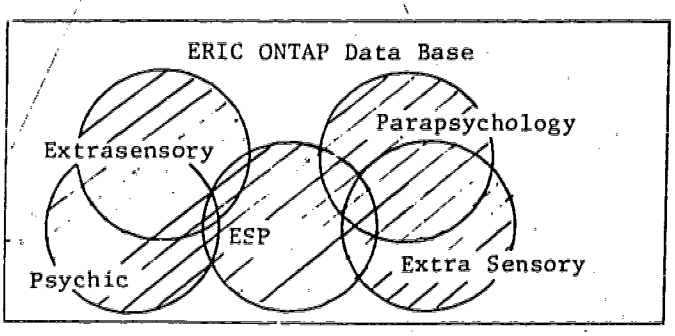
There is no mention concerning limiting the output by publication type, availability, etc. When selecting an approach to search strategy, the searcher must take into consideration the search objective, to retrieve all information on this topic. In this case, a high recall formulation is preferred; moreover, the searcher should make every effort to compile a comprehensive list of search terms in order to meet the search objective.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Search strategy is illustrated in the following diagram. The rectangle contains all records stored in the ERIC ONTAP data base. A shaded circle is used to represent the set of documents in which the concept parapsychology occurs.



In the next diagram, the shaded circles indicate not only the sets of documents formed by search terms, but the retrieved portion of the data base. The illustration depicts formulation #1 in which the controlled vocabulary fields are scanned for the occurrence of the search terms.



Searching the assigned index term fields is limited to the identifier field since there are no ERIC descriptors to represent the concept. Since identifier words and phrases describe the intellectual content of the document, this search is expected to retrieve a small number of non-relevant items, thus it is a high precision formulation.

Formulation #1			Comments
1	S	PARAPSYCHOLOGY/ID	
2	S	ESP/ID	
3	S	EXTRA SENSORY PERCEPTION/ID	Two different spellings of extra sensory are entered in order to retrieve variant forms
4	S	EXTRASENSORY/ID	
5	S	PSYCHIC/ID	
6	C	1-5/OR	Abbreviated form of COMBINE 1 OR 2 OR 3 ...

Since the search objective is to retrieve all information on this topic, Formulation #2 is constructed to search all other fields containing subject describing words and phrases. This formulation employs the same terms as the previous search but is not restricted to searching only the assigned index term fields. The strategy utilizes the "free text searching" of the DIALOG retrieval system in which all fields containing subject terms are scanned: title, abstract, identifier, corporate source, sponsoring agency, descriptor, and descriptive note. In addition, word proximity and truncation features are incorporated into the strategy in order to retrieve word variants.

Formulation #2

1 S PARAPSYCH?

2 S ESP

3 S EXTRA(W)SENSORY

4 S EXTRASENSORY

5 S PSYCHIC?

6 C 1-5/OR

Both formulations are examples of the building block approach to search strategy. This approach is especially suitable for this topic; since all information on this topic is desired, the searcher need not be concerned with restricting

the size nor with developing a high precision formulation. Consequently, the searcher is more involved with expressing all aspects of the concept, word variants and synonyms, rather than monitoring intermediary results of the search. The Briefsearch would appear unnecessary here since the topic is limited to a single facet and can be easily searched using a few simple search statements.

(Reader: Keeping in mind the stated information need of the requestor, i.e., to retrieve all relevant items in the ERIC ONTAP file, which formulation would you choose? _____)

Before evaluating the results of the two formulations, you may wish to draw up your own strategy and search the ONTAP file. Formulation #2 serves as an example for demonstrating the searching procedures of the ONTAP file. Whether you construct your own formulation or use one of the above examples, follow the explanation below as you use the ONTAP file.

After signing on to DIALOG, enter the BEGIN 201 command to start searching the file. Let's review the explanation of the ONTAP file. Type ?ONTAP to view the text.

```
ONTAP(ONLINE TRAINING & PRACTICE) FILE
PROVIDES TEST QUESTIONS & RELEVANT
CITATIONS FOR EACH QUESTION.  THE 1975
ERIC FILE IS THE SOURCE OF THE ANSWER
SETS.  QUESTIONS AND ANSWER SETS ARE
LABELLED AS TO COMPLEXITY,E.G. AN=S03,
AN=M04,AN=D03 DENOTE THE ANSWER SETS
FOR THE 3RD SIMPLE, 4TH MEDIUM, AND 3RD
DIFFICULT QUESTION.  AFTER FINISHING THE
SEARCH, THE PERFECT ANSWER SET FOR THAT
QUESTION CAN BE SELECTED BY COMMANDS OF
THE FORM 'SELECT AN=S03', AND SELECTED
ANSWER SETS CAN BE COMBINED IN AN 'AND'
RELATIONSHIP WITH YOUR FINAL SEARCH SET
TO FIND OUT HOW MANY RELEVANT CITATIONS
YOU FOUND.  A COMMAND /EVAL COMPUTES THE
RECALL AND PRECISION OF YOUR SEARCH.
SEE   ?TAPSAM FOR SAMPLE SIMPLE SEARCH
      ?TAPSIM FOR LIST OF SIMPLE QUESTS.
      ?TAPMED FOR LIST OF MEDIUM QUESTS.
      ?TAPDIF FOR LIST OF DIFFICULT
       QUESTIONS
      ?EVAL FOR EXPLANATION OF PREC/REL.
```

At the conclusion of the above explanation are listed a number of commands; the command, ?TAPSAM, provides step-by-step directions for searching the ONTAP file. ?TAPSIM, ?TAPMED, ?TAPMED2, ?TAPDIF, ?TAPDIF2, are lists of ONTAP search questions and their corresponding "answer set." The command, ?EVAL, gives the procedures for computing recall and precision scores. Using formulation #2 as an example, we are ready to perform the search. Let's review all the steps of the searching procedures by reading the directions triggered by the ?TAPSAM command.

```
? ?TAPSAM
TO USE ONTAP FILE, THE USER MUST BEGIN201:
  1)SELECT QUESTION FROM PRINTED LIST OR
  BY ?TAPSIM, ?TAPMED, ?TAPDIF,
  2)COMPLETE SEARCH IN YOUR OWN WAY
  3)SELECT AN=(YOUR CHOSEN QUESTION #) FOR
  ANSWER SET
  4)AND YOUR FINAL SEARCH SET AND ANSWER
  SET
  5)ENTER EVAL COMMAND USING /EVAL(A,B,C)
  WHERE A=# OF CITATIONS IN SEARCH SET,
  B=# OF CITATIONS IN ANSWER SET, C=# OF
  CITATIONS IN COMBINED SET
E.G.: 1)CHOOSE QUESTION AN=S01
      2)S DRAW(2W)PERSON(W)TEST (SET 1)
      3)SELECT AN=S01 (SET 2)
      4)COMBINE 1 AND 2
      5)ENTER /EVAL COMMAND,E.G. /EVAL
      (1,2,1)
TO IMPROVE YOUR RESULTS, TRY LOOKING AT
OTHER RECORDS FOR SEARCHING IDEAS.
```

The first step in the directions recommends that we select a search topic from the lists of simple, medium, and difficult questions. Since we have already chosen parapsychology as a topic, let's move on to the second step, "complete search in your own way." Formulation #2 is repeated below, but you can perform the search using your own strategy.

Formulation #2

```
S  PARAPSYCH?
     1    6  PARAPSYCH?
S  ESP
     2    7  ESP
S  EXTRA(W)SENSORY
     3    2  EXTRA(W)SENSORY
S  EXTRASENSORY
     4    2  EXTRASENSORY
S  PSYCHIC?
     5    9  PSYCHIC?
C  1-5/OR
     6   18  1-5/OR
```

At this point, it is necessary to find out whether parapsychology is a simple, medium, or difficult topic so that we can compare our answer set with ONTAP's "perfect answer set." Since there are a number of search question lists, let's print the list of simple questions first using the ?TAPSIM command.

```
? ?TAPSIM
SIMPLE SEARCH TOPICS:
S01  DRAW-A-PERSON TEST
S02  FID(INT. FED. FOR DOC.)&LIBRARY
OR INFORMATION NETWORKING
S03  4-H CLUBS,THEIR MEMBERS &ACTIVITIES
S04  REVISION,ANGLO-AMERICAN CATAL.RULES
S05  NAVAHO LANGUAGE TEXTBOOKS/GRAMMARS
(MATERIAL IN NAVAHO,OR USEFUL FOR
TEACHING NAVAHO,OR ABOUT NAVAHO LINGUIS-
TICS)
S06  EDUCATION,SRI LANKA(INCL.LIBRARY
ACTIVITIES)
S07  LITERACY IN DEVELOPING COUNTRIES
PUBLISHED BY THE INT'L. INSTITUTE FOR
ADULT LITERACY METHODS (OR ANY PUBLICA
TIONS BY OR ABOUT THIS INSTITUTE)
S08  16 PERSONALITY FACTOR TEST
S09  PARAPSYCHOLOGY
```

Reviewing the list above, we find that parapsychology is simple search question number 9. By selecting the accession number assigned to the topic, we can obtain the number of postings in the "perfect answer set." The following example demonstrates how to select the answer set.

Formulation #2 (cont.)	Comments
SELECT AN=SO9 (or S AN=SO9)	Selecting answer set to simple question #9
7 9 AN=SO9 (Parapsychology)	System response

The next step is to combine our final answer set (search statement #6) with the ONTAP answer set (search statement #7). In this way, we can find out how many items we retrieved are included in the ONTAP answer set. The result of this operation is shown below, and indicates that we have retrieved all the items in the ONTAP answer set in addition to some non-relevant material.

Formulation #2 (cont.)	Comments
C 6 AND 7	Statements 6 and 7 are combined excluding non-relevant material in our retrieved set
8 9 6 AND 7	System response

Although we can figure recall and precision scores by comparing the relevant items listed by ERIC accession number at the conclusion of this discussion with the final set (statement #6) of our search, ERIC ONTAP provides the /EVAL capability for computing the scores. The directions are obtained by typing ?EVAL.

```
? ?EVAL
/EVAL(A,B,C) COMPUTES THE RECALL AND
PRECISION OF A SEARCH.  THE PARAMETERS
ARE:
     A=TOTAL NUMBER OF CITATIONS FOUND
       IN YOUR FINAL ANSWER SET
     B=NUMBER OF CITATIONS IN ONTAP
       ANSWER SET
     C=NO. OF CITATIONS IN ANSWER SET
         FOUND BY USER (RESULT OF
LOGICAL AND OF A AND B)
   RECALL=C/B   PRECISION=C/A
RECALL TELLS YOU % OF RELEVANT CITA-
TIONS IN THE FILE THAT YOU FOUND
PRECISION GIVES % OF YOUR CITATIONS
THAT WERE RELEVANT
```

After following the directions above, we can determine from reviewing the online interaction that A = 18 (total number of citations found in my final answer set, statement #6), B = 9 (number of citations in ONTAP answer set, statement #7),

and C = 9 (number of citations in answer set found by COMBINing statements #6 and #7, statement 8). These values are inserted into the /EVAL formula in order to obtain recall and precision percentages.

Formulation #2 (cont.)	Comments
/EVAL (18,9,9)	EVAL command with values A, B & C inserted
INPUT...EVAL(18,9,9)	System response to /EVAL formula

Number of citations found = 18

Number of citations in answer set = 9

Number of relevant citations found = 9

RECALL = 100% PRECISION = 50%

As mentioned previously, you can figure recall and precision ratios for your performance by checking the results of your search with the list of ERIC accession numbers which form the answer set to the parapsychology search topic. Before moving on to the next search question, let's compare the results of the two formulations for parapsychology.

Step 8: Evaluating Final Results

As shown above, the second formulation which uses free text searching of the ONTAP data base retrieves all the relevant citations. The results of the search are summarized below:

Formulation #2

Number of citations found = 18

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{9}{9}$$ (100% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{9}{18}$$ (50% of retrieved documents are relevant)

Of the eighteen citations retrieved by formulation #2, nine are non-relevant; seven of these are retrieved by the search term psychic used as a homograph in the title or abstract. "Psychic energy," "psychic distance," and "psychic development," serve as examples of phrases in which the term psychic occurs

but is deemed non-relevant to the search topic. In the remaining two non-relevant items, ESP is used as an acronym for a program.

In contrast, formulation #1 retrieves six citations. Comparing its answer set with the list of ERIC accession numbers at the conclusion of this discussion, we can obtain recall and precision.

Formulation #1

Number of citations found = 6

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{6}{9} \quad \text{(67\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{5}{6} \quad \text{(83\% of retrieved documents are relevant)}$$

This formulation relies on the controlled vocabulary for search terms; in this case, precision, the number of relevant items in the output, improved when compared with the precision of the second formulation. The one non-relevant item is retrieved because ESP occurs as an acronym in the identifier field. Of the four relevant items which are not retrieved, two contain ESP as a term in the title and two contain parapsychology as a term in the abstract. The formulation fails to capture these four relevant documents because both the title and abstract fields, two sources of subject-rich terminology, are not searched.

Since the requestor is interested in retrieving all relevant information, formulation #2 is preferred in this instance.

Relevant citations (cited by ERIC accession number):

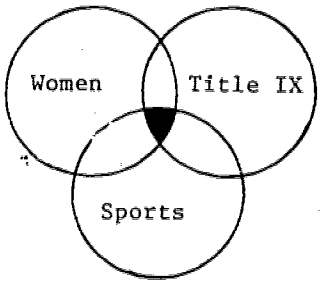
EJ115774 EJ113605 EJ113604 EJ110046 EJ105855 ED110382
ED107626 ED104113 ED099252

Search Topic #1: Title IX, 1972 Federal Education Act Amendments, and Women's Sports. (S AN=M10)

Search Objective: The expressed information need is to retrieve high recall with a low number of non-relevant citations.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

Three facets comprise the topic: sports, women, and Title IX. The diagram below illustrates the three facets of the search. Where the three circles intersect is the retrieved portion of the data base.



Subsearches for the three facets of the search assume the same strategy as the sample search topic parapsychology. Just as all terms representing the concept of parapsychology are searched and the results combined in an OR relationship, subsearches for each facet of this topic will be performed in the same way. All terms representing WOMEN, free text and/or assigned index terms are SELECTed and then COMBINEd by an OR operator; likewise, SPORTS, and its variant forms and synonymous terminology are joined by OR; a set representing TITLE IX is formed similarly. The three subsearches are then COMBINEd by the AND operator, thus creating the final set.

Step 4: Compiling the Search Terms

Women is a "lead-in" term in the ERIC Thesaurus directing the searcher to females. From the list for females, descriptors indicating WOMEN are gathered. The term sports like women, is not a descriptor but leads to Athletics. There are no descriptors to represent Title IX or 1972 Federal Education Act Amendments; consequently, determining variant forms of the Act for searching the assigned index fields of the ERIC record can be done by reviewing the online dictionary for different forms which may occur in the identifier field, SELECTing from the list, and/or using these forms to compose word proximity phrases for free text searching.

Step 5: Ordering Output

Since the information seeker desires a high recall, with a minimum of non-relevant citations in the output, the searcher must decide what approach to search strategy would eliminate the majority of non-relevant citations from the output. Selecting the most specific facet first approach may be a wise choice as the size of the set representing the most specific facet, TITLE IX, may be small enough to terminate the search before applying any other facets to the strategy. At the same time, the output can be monitored for loss of relevant material as search facets are applied to the strategy. The Briefsearch could also be used initially in order to estimate the size of the final set. The formulations below include both approaches as well as the building block approach.

Step 6: Conceptualizing the Search as Input to the Retrieval System

The Briefsearch, formulation #1, can be used to make a quick survey of the ERIC file before a more complicated formulation is entered; rapid and efficient in terms of online connect time, the Briefsearch can be formulated in a number of ways.

Briefsearch

1 S WOMEN(C)ATHLETICS(C)TITLE(W)IX

2 S WOMEN'S(C)ATHLETICS(C)TITLE(W)IX

3 C 1 OR 2

Comments

Use of the (C) as the logical operator AND eliminates additional steps for SELECT and COMBINE

The Briefsearch formulation retrieves three citations. At this point, the searcher may call up the citations for online printing in order to check the searching vocabulary used in the strategy with that of the free text and assigned index fields of the document, or the searcher may continue searching with a more complex formulation.

In formulation #2, all subject-conveying fields are searched using free text and word proximity. Since WOMEN, one of the facets of the topic, is a concept which occurs frequently as a population group in ERIC searches, a subsearch is formulated with high recall as the goal; the search is saved for future use in ERIC searches. By entering the .EXECUTE command, the Search Save for WOMEN is performed and its results can be assimilated into the present search for Title IX and Women's Sports. The formulation of the Search Save is given below. The appendix contains 12 Search Save formulations, which are referred to in Section II search topics.

Search facet: Women

1 S DAUGHTER?
2 S FEMALE?
3 S FEMINI?
4 S GIRL?
5 S GRANDMOTHER?
6 S HOUSEWI?E?
7 S LADY?
8 S LADIES?
9 S MOTHER?
10 S NUN

Comments

All subject-conveying fields are searched in the Search Save for WOMEN. High recall is the primary goal of the subsearch.

Included DESCRIPTORS and identifiers are FEMALES, FEMINISM, GIRLS CLUBS, HOUSEWIVES, MOTHER ATTITUDES, MOTHERS, NUN TEACHERS, NUNS, WOMEN PROFESSORS, WOMEN TEACHERS, WORKING WOMEN, WOMEN'S ATHLETICS, WOMEN'S EDUCATION, WOMEN'S STUDIES/Female, Feminist Press, Feminization, Masculinity Feminity Variable, Girl Guides Association, Girl Scouts, Girl Scouts of America

Search facet (con't): Women

```
11 S NUNS
12 S SISTER
13 S SISTERS
14 S WIFE
15 S WIVES
16 S WOM?N?
17 C 1-16/OR
   END/SAVE
SERIAL#1KLM
      23MAY78 8:52:56 USER 4111
$0.00  0.046  HRS FILE201  00 DESCRIPTORS
$0.37  TYMNET
$0.37  ESTIMATED TOTAL COST
```

Comments

Included DESCRIPTORS and identifiers are Camp Fire Girls, 3 multi-term MOTHERS Descriptors, 7 multi-term Mother identifiers, Sisters, Wife, Wives, Women, 11 multi-term Women and Women's identifiers

Saved by the END/SAVE command, the formulation is assigned serial number 1KLM. The Search Save can be recalled and executed by the command .EXECUTE 1KLM
Procedures for exchanging Search Saves between DIALOG accounts is given in the introduction to the appendix.

Free text searching is used in formulation #2. The building block approach characterizes this strategy as all three facets are developed in the formulation and combined in the final step.

Formulation #2: Free text searching

```
1 .EXECUTE 1KLM
2 S ATHLET?
3 S GYM?
4 S PHYSICAL(W)EDUCATION
5 S SPORT?
6 C 2-5/OR
7 S EDUCATION(W)AMENDMENTS(F)1972
8 S EDUCATION(W)AMENDMENT(F)1972
9 S TITLE(W)9
```

Comments

Search Save for WOMEN is processed
Included DESCRIPTORS/Identifiers: ATHLETES, ATHLETIC COACHES, ATHLETIC EQUIPMENT, ATHLETIC FIELDS, ATHLETIC PROGRAMS, INTRAMURAL ATHLETIC PROGRAMS, EXTRAMURAL ATHLETIC PROGRAMS, ATHLETICS, WOMEN'S ATHLETICS/Athletic Administrators, Athletic Associations, Athletics for Handicapped, Athletic Trainers
GYMNASIUMS/Gymnastics, Gymnastics Judges
PHYSICAL EDUCATION, ADAPTED PHYSICAL EDUCATION, PHYSICAL EDUCATION FACILITIES/ Physical Education and Recreation, Physical Education Majors, Physical Education Teachers
SPORTSMANSHIP/Sport Ambassadors, Sport for All, Sporting Goods, Sports, 6 multi-term Sports Identifiers

10 S TITLE(W)IX — Title IX

11 C 7-10/OR

12 C 1 AND 6 AND 11 — All three facets are joined together in this final statement

Working with the most specific facet first approach allows the searcher to terminate the search when the final set shrinks to a size satisfactory to the search objective. TITLE IX is selected as the most specific facet since it is a named object rather than a concept like WOMEN or SPORTS. In order to determine the different ways the phrase can be represented, the searcher uses the EXPAND capability, SELECTs from the display, and formulates word proximit phrases for free text searching. The EXPAND display for the phrase TITLE IX/ID follows:

REF	INDEX-TERM	TYPE	ITEMS	RT
E1	TITLE I----------		2	
E2	TITLE I HEA----------		1	
E3	TITLE I PRESCHOOL PROGRAMS----------		1	
E4	TITLE III----------		1	
E5	TITLE IV----------		1	
E6	-TITLE IX----------		12	
E7	TITLE IX OF THE EDUCATIO N AMENDMENTS ACT O----		1	
E8	TITLE V----------		1	
E9	TITLE VII----------		2	

From the above display, the searcher can either select the desired items or go on to another EXPAND display for the Education Amendments Act of 1972.

E3	EDUCATION----------		22153115
E4	EDUCATION ACT OF 1972---		1
E5	EDUCATION AMENDEMENTS OF 1974----------		1
E6	-EDUCATION AMENDMENTS----		1
E7	EDUCATION AMENDMENTS OF 1972----------		1
E8	EDUCATION AMENDMENTS OF 1972 TITLE IX---------		1
E9	EDUCATION AMENDMENTS OF 1974----------		1
E10	EDUCATION AMENDMENTS 1972----------		1
E11	EDUCATION AMENDMENTS 1972 TITLE IX---------		5
E12	EDUCATION AMENDMENTS 1974----------		1
E13	EDUCATION AND ECSTASY---		1

After viewing the EXPAND displays, the searcher can express the TITLE IX facet in word proximity phrases.

Formulation #3	Comments
1 S TITLE(W)IX	
2 S TITLE(W)9	
3 S EDUCATION(W)AMENDMENTS(F)1972	
4 S EDUCATION(W)AMENDMENT(F)1972	
5 C 1-4/OR	At this point, 39 citations are retrieved. Rather than end the search, one of the other facets is applied to the formulation
6 S ATHLET?	
7 S SPORT?	
8 C 5 AND (6 OR 7)	The combination of the two facets, TITLE IX and SPORTS results in 8 citations. The search could be ended or the third facet applied

In the above formulation, the application of the second facet to the first reduces the postings from 39 to 8. The searcher could develop the third facet but, in this case, has terminated the search rather than continue the online interaction; however, this choice may affect the precision score as all facets representing the topic have not been exhausted in the strategy.

(Reader: In the light of the search objective, to retrieve high recall with a low number of non-relevant citations, which strategy would you choose? _____)

Step 8: Evaluating Final Results

Before figuring recall and precision scores, we must find out what accession number has been assigned to the search in the ONTAP file. Reviewing the list of all simple topics in the discussion of the parapsychology query discloses that this topic on Title IX and Women's Sports is not among the simple topics. The list of medium topics may be obtained by the command ?TAPMED.

```
? ?TAPMED
MEDIUM SEARCH TOPICS:
M01  DIRECT CHARGING TO USERS FOR REF-
EREECE & CURRENT AWARENESS SERVICE OF
     LIBRARIES OR OTHER INFO SERVICES(
     PHILOSOPHY,POLICY,PRACTICE, FEE
     CHARGES;ANY TYPE OF LIBRARY;ANY
     TYPE OF REFERENCE SERVICE;NOT IN-
     TERESTED IN FREE SERVICES)
M02  FEDERAL AID TO DAY CARE CENTERS OR
     SERVICES (INCLUDING HISTORY,PHILOS-
     OPHY,ARGUMENTS PRO & CON,EXPERI-
     ENCES,FUNDING,EVALUATION,PARENT
     INVOLVEMENT AND ATTITUDES) ONLY
     U.S. CENTERS OR SERVICES
M03  JEAN PIAGET'S THEORIES, AND THOUGHT
     PROCESSES OR LANGUAGE DEVELOPMENT
     OF CHILDREN,BUT LIMITED TO PUBLI-
     CATION AVAILABLE FROM ERIC/EDRS
FOR MORE, ENTER ?TAPMED1
? ?TAPMED1
MEDIUM SEARCH TOPICS:
M04  LIBRARIES & LIBRARIANS IN MIDDLE
     EAST,(EXCLUDE AFRICAN BUT NOT
     EGYPT).  INCLUDE ALL TYPES OF LIB-
     RARIES & INFORMATION CENTERS
M05  LIBRARY SERVICE TO PHYSICALLY
     HANDICAPPED (NOT MENTALLY OR
     LANGUAGE-HANDICAPPED)
M06  EFFECTS OF TV VIOLENCE ON CHILDREN
M07  DRUG ABUSE,INCLUDING ALCOHOL,AMONG
     STUDENTS OF ELEMENTARY OR SECONDARY
     SCHOOLS,GRADES K-12(INCLUDING
     SCHOOL EDUCATION PROGRAMS AND
     SOCIOLOGICAL STUDIES)
M08  SCHOOL BUSING & RACIAL INTEGRATION
M09  RECREATIONAL USE OF FOREST LANDS
M10  TITLE 9,1972 FEDERAL EDUCATION ACT
     AMENDMENTS,AND WOMEN'S SPORTS
M11  WHITE FLIGHT TO THE SUBURBS
```

Scanning the above list, one finds that the search question has been assigned accession number M10. To determine how many citations are relevant to the search topic, select the accession number, e.g., SELECT AN = M10. Using the EVAL procedure, discussed in the sample search, the results of the three formulations are evaluated. The Briefsearch, formulation #1, fared well, retrieving three citations, two of which are relevant; in this case, it functions as a survey of the file, retrieving 50% of the relevant items in the file.

Briefsearch

Number of citations found = 3

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{2}{4}$ (50% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{2}{3}$ (66% of retrieved documents are relevant)

Formulation #3 in which the most specific facet first approach is demonstrated retrieves all the relevant items but also as many non-relevant items.

Formulation #3

Number of citations found = 8

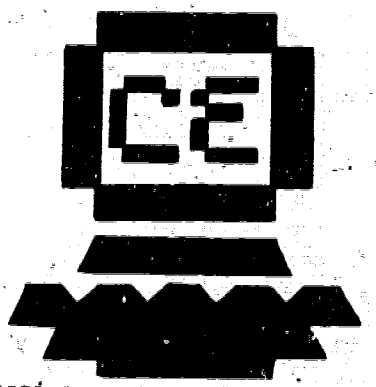
Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{4}{4}$ (100% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{4}{8}$ (50% of retrieved documents are relevant)

Formulation #2 retrieves all the relevant items, no non-relevant items, resulting in 100% recall and 100% precision. As the information need is stated, the requestor would be satisfied with a high recall search as long as few non-relevant citations are included in the output. Both formulations #2 and #3 perform well in terms of recall; however, the precision of formulation #3 in which the most specific facet first approach is used, is lower than formulation #2, the building block approach. Formulation #3 is more efficient and rapid in terms of online connect time since only two facets are developed in the strategy. Thus, either formulation #2 or #3 is preferred for the search topic.

Relevant Citations:

EJ119209 EJ105832 EJ105671 ED110452



Search topic #2: 4-H Clubs, their members and activities.

(S AN=S03)

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

A single concept is present in the topic: 4-H Clubs. Consideration must be given when constructing the formulation to include variant forms and spellings of the Club.

Step 4: Compiling the Search Terms

No descriptors are listed in the ERIC Thesaurus which directly refer to 4-H Clubs. Search formulations are constructed using the word proximity feature in order to retrieve different spellings of the Club.

Step 5: Ordering Output

No output specifications are requested.

Step 6: Conceptualizing the Search as Input to the Retrieval System

If a high precision search is required, you could search only the identifier field as in the following formulation:

Formulation #1	Comments
1 S 4(W)H/ID	Using the word proximity feature, additional citations are retrieved in cases where another word. e.g. 4 H Clubs, follows the identifier phrase. Without the proximity feature, only two citations are retrieved in comparison to 14 citations
2 S FOUR(W)H/ID	
3 C 1 OR 2	

The following search formulation is a high recall formulation in which all subject-conveying fields are searched:

Formulation #2	Comments
1 S 4(W)H	All subject-conveying fields are searched in this formulation
2 S FOUR(W)H	
3 C 1 OR 2	

Step 7: Evaluating Preliminary Results

In formulation #1 fourteen citations are retrieved, 21 in the second formulation. At this point the number of citations is sufficiently small in both cases that it is felt to be unnecessary to impose the "member" or "activities" aspect of the search. This is an example of starting with the most specific facet first.

Step 8: Evaluating Final Results

Searching the identifier field in the first formulation results in high precision but some relevant items are missed resulting in a lower recall score than formulation #2 in which all subject rich fields are searched. The recall and precision scores for both formulations are summarized below:

Formulation #1

Number of citations found = 14

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{14}{17}$$ (82% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{14}{14}$$ (100% of retrieved documents are relevant)

Formulation #2

Number of citations found = 21

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{17}{17}$$ (100% relevant documents retrieved)

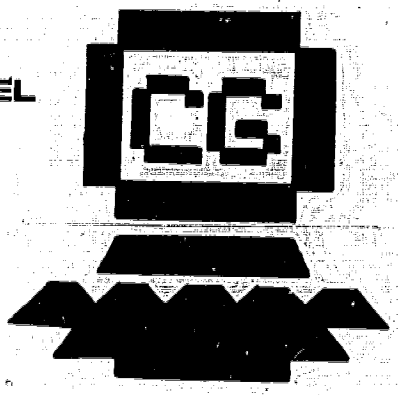
$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{17}{21}$$ (85% of retrieved documents are relevant)

Relevant citations:

EJ120500 EJ117540 EJ113715 EJ109322 EJ107077 EJ107045 ED109352

ED109351 ED109205 ED105103 ED105102 ED105101 ED105100 ED104895

ED102469 ED100323 ED095289



Search Topic #3: 16 Personality Factor Test

Search Objective: The stated information need is to retrieve a maximum number of relevant citations using the least amount of online connect time. (S AN=S08)

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

A single facet is present in the search topic: 16 Personality Factor Test. The main difficulty with this search is the different spellings of this test.

Step 4: Compiling the Search Terms

Since no descriptors are listed in the ERIC Thesaurus, you could EXPAND the numeric character 16, as an identifier, SELECT from the online dictionary; then EXPAND 16 spelled-out and SELECT from that portion of the online dictionary for identifiers.

Step 5: Ordering Output

Keeping in mind the search objective, the searcher should strive to reduce online connect time by preparing before the online interaction, curtailing excessive typing in order to minimize typing errors, and selecting a suitable approach to search strategy. The Briefsearch may be favored because of its efficiency; variant spellings of the test could be represented by single search statements using word proximity.

Step 6: Conceptualizing the Search as Input to the Retrieval System

The following formulation takes into account different spellings of the test garnered from online dictionary displays:

Formulation #1

1 S 16PF
2 S 16(W)PF
3 S 16(W)P.(W)F.
4 S 16(W)PERSONALITY(W)FACTOR
5 S 16(W)PERSONALITY(W)FACTORS
6 S SIXTEEN(W)PERSONALITY(W)FACTOR
7 S SIXTEEN(W)PERSONALITY(W)FACTORS
8 C 1-7/OR

Comments

Free text searching is used exclusively in this formulation.

Includes Identifiers: 16PF, 16 P. F., 16 Personality Factor Questionnaire, 16 Personality Factor Test, Sixteen Personality Factor Questionnaire

As an abbreviated form of the above formulation, the following Briefsearch contains all the elements of the longer search except search statements 2 and 3. Since two possible ways of spelling the test are not included in the formulation, a lower recall score may be expected.

Briefsearch

1 S 16PF
2 S 16(W)PERSONALITY
3 S SIXTEEN(W)PERSONALITY
4 C 1-3/OR

Comments

Includes all of the Identifiers named above except 16 PF and 16 P. F.

(Reader: Keeping in mind that the information seeker is concerned with online connect time, which formulation would you choose? _____)

Step 8: Evaluating Final Results

In order to find out what citations are relevant to the topic, the searcher must review the lists of simple, medium, and difficult topics to found out the accession number of the search. From the list of simple topics, we find that the topic is assigned number simple 8 (S08) which we then SELECT (S AN=S08) in order to obtain the number of relevant items. Surprisingly, both formulations

fared the same, retrieving all 12 relevant items and one non-relevant citation. Recall and precision scores for both formulations are summarized below:

Formulation #1 and Briefsearch

Number of citations found = 13

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{12}{12} \quad \text{(100\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{12}{13} \quad \text{(92\% of retrieved documents are relevant)}$$

Since the online dictionary is called up in formulation #1, it is time-consuming, thus costly. In a trial run of the first formulation in which four EXPAND commands are used, the search takes .125 (7½ minutes) to perform. In comparison, the Briefsearch takes .25 (1½ minutes), and, in this case, retrieves the same set of documents. Because of its efficiency, the Briefsearch is preferred for search topic #3.

Relevant citations:

EJ120592 EJ117340 EJ117192 EJ115601 EJ115600 EJ112211

EJ108963 ED103575 ED103574 ED102449 ED101228 ED095465

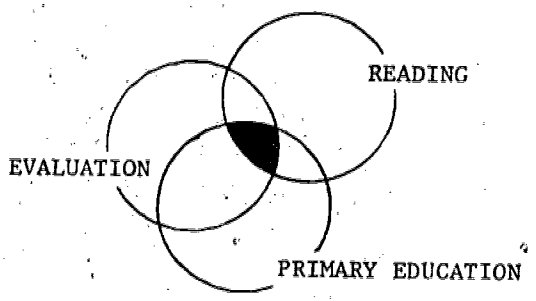
READING AND COMMUNICATION SKILLS

Search Topic #4: Evaluation of Primary School (grade K-3) English reading programs or reading materials and techniques (not evaluation of specific reading tests or instructors, not student test scores when not part of evaluation of reading program, and not just criteria or standards for evaluation). (S AN=D03)

Search Objective: To retrieve as much relevant material as possible with a minimum of non-relevant citations in the output. Only publications available from EDRS are desired.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

The search topic can be broken down into three facets: reading, primary school, and evaluation. Illustrated in the following diagram, the search topic is separated into its three facets which intersect to form the portion of the data base intended for retrieval.



Representing those aspects of the query which are not wanted, e.g., evaluation of reading tests, evaluation of instructors, may eliminate relevant material from the output. For example, a set created by the statement "EVALUATION AND READING" may have documents dealing with the evaluation of reading teachers as well as how teachers evaluate reading materials, so that a set created by the statement

"READING AND (EVALUATION NOT TEACHERS)" would exclude valuable material. Consequently, the search formulations constructed below will not employ the Boolean NOT operator.

Step 4: Compiling the Search Terms

All facets of the search are well represented by terms from the ERIC Thesaurus. In fact, entering READING as a free text search term retrieves occurrences of the term in 58 multi-term READING descriptors, 52 multi-term READING identifiers and all occurrences in free text fields, e.g. title, abstract or corporate source. EVALUATION occurs in 26 multi-term descriptors, and in 10 multi-term identifiers as well as in free text fields. The abundance of subject terminology acts detrimentally as the search is made difficult by the volume of literature in the ERIC file on reading and on evaluation.

Step 5: Ordering Output

In order to restrict the search to documents available from EDRS, the LIMIT command can be applied at the beginning of the search. Applying the LIMIT after the creation of the first facet allows the searcher to monitor preliminary search results throughout the online interaction.

The ERIC file contains much material on reading and evaluation, so that the searcher should select an approach to search strategy which enables constant evaluation of preliminary search results. In this way, the searcher can introduce limiting criteria to the formulation as the search is in progress. For example, in the building block approach, the combination of all three facets occurs as the final statement of the search; if the resultant set is too large, the searcher has little recourse except to "backtrack." or reformulate some of the essential elements of the search.

On the other hand, in the successive fractions approach, the

search is limited to the initial "bite" or partition of the file; subsequent search facets continue to be applied to this partition so that the searcher can terminate the search if necessary before exhausting all search facets.

The sample formulations provided in step 6 cover the successive fractions and building block approaches. Follow the development of the formulations especially by reading the COMMENTS section, so that you are aware of the decisions made by the searcher as the search proceeds.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Because the volume of literature on Reading is so large, you could formulate a Briefsearch first in order to perform a survey of the file. This will give you an idea of the possible size of the final set.

Briefsearch	Comments
1 S PRIMARY(W)EDUCATION(F)READING(F)EVALUATION	Twenty citations are retrieved in the Briefsearch giving an estimation of the volume of output
2 S PRIMARY(W)GRADES(F)READING(F)EVALUATION	
3 C 1 OR 2	
4 LIMIT 3/AVAIL	

Since the terms READING and EVALUATION have such high posting counts, the formulation might concentrate on the identifier, descriptor, and title fields in order to eliminate the possibility of retrieving false drops due to casual mention of the search terms Reading and Evaluation in the corporate source, abstract or descriptive note fields. Using Search Save #1 (see appendix) for representing the facet of PRIMARY EDUCATION, the following strategy, an example of the building block approach, searches the assigned index fields and title for the READING and EVALUATION facets and all free text fields for the PRIMARY EDUCATION facet.

Formulation #2	Comments
1 .EXECUTE____(Search Save serial number for Primary Education)	
2 LIMIT 1/AVAIL	The LIMIT is applied to the set of the facet PRIMARY EDUCATION as the second step in this formulation
3 S READING/DE,ID,TI	Reading, as a search term, is restricted to occurrences in the title and assigned term fields.
4 S COMPARATIVE(W)ANALYSIS/DE,ID,TI	Statements 4-8 form the facet EVALUATION; searching is limited to the title and assigned term fields.
5 S COMPARATIVE(W)STATISTICS/DE,ID,TI	
6 S COMPARATIVE(W)TESTING/DE,ID,TI	
7 S EVALUATION/DE,ID,TI	
8 C 4-7/OR	
9 C 2 AND 3 AND 8	

The high recall search formulation employs Search Save #1 for the PRIMARY EDUCATION facet and Search Save #7 for the EVALUATION facet (see appendix). Formulation #3 is an example of the successive fractions approach to search strategy; as this search progresses, note what factors are applied to the formulation in order to restrict volume of output.

Formulation #3	Comments
Part I:	
1 .EXECUTE____(Search Save for Primary Education)	
2 LIMIT 1/AVAIL	
3 S READER?	
4 S READING	
5 C 2 AND (3 OR 4)	Two facets of the search, READING and PRIMARY EDUCATION are joined by AND; the large number of citations (307) retrieved at this point makes it clear that further reductions are necessary. The next step is to apply the general Search Save for EVALUATION to the search.
6 .EXECUTE____ (Search Save for general EVALUATION)	

Formulation #3 (cont.)

Comments

Part I (cont.):

7 C 5 AND 6

The COMBINing of the three facets results in 250 citations. At this point the search could be concluded if the search objective is high recall.

Part II:

8 .EXECUTE_____ (Search Save for restricted EVALUATION)

Since the output of the high recall formulation is large, the formulation is reconstructed using Search Save #8 for EVALUATION. Search terms are limited to descriptors whose occurrences are retrieved from the assigned index and title fields.

9 C 8 AND 5

Here, the volume of output is reduced to a total of 121 citations

10 S READING/DE,ID,TI

Another approach to restricting the formulation is to limit the READING facet to occurrences of Reading in title, identifier and descriptor fields.

11 C 2 AND 6 AND 10

At this point, 196 citations are retrieved, and it is decided to continue introducing limiting criteria to the search formulation.

Part III:

12 S ACHIEVE?

Instead of using the comprehensive EVALUATION Search Saves, a formulation containing a few search terms from the EVALUATION Search Saves is constructed.

13 S APPRAIS?

Statements 12-17 make up a subsearch for EVALUATION

14 S ASSESS?

15 S COMPAR?

16 S EVALUAT?

17 C 12-16/OR

18 C 2 AND 10 AND 17

137 citations are retrieved using the abbreviated subsearch for EVALUATION, Search Save for PRIMARY EDUCATION, and the restricted READING set.

(Reader: Given the many possible approaches to formulating the search strategy, which formulation would you choose? Keep in mind that the information seeker desires a minimum of non-relevant citations in the output._____)

Step 8: Evaluating Final Results

The list of difficult questions in the ONTAP file is given below and must be checked for ONTAP accession numbers. In this way, one can SELECT the accession number (S AN=D03) in order to find out how many citations are relevant to this topic.

```
? ?TAPDIF
D01  TRAINING FOR SUPERVISION & MANAGE-
     MENT IN LIBRARIES & INFORMATION
     CENTERS(INCL. NEED FOR TRAINING,DES
     CRIPTIONS OF TRAINING PROGRAMS OR
     MATERIALS,TRAINING OF STUDENTS AND
     PROFESSIONAL WORKING LIBRARIANS;
     ACADEMIC OR ON-THE-JOB TRAINING
D02  AUDIOVISUAL AIDS FOR
     ORIENTATION OR INSTRUCTION OF
     LIBRARY USERS
D03  EVALUATION OF PRIMARY SCHOOLS(GRADE
     K-3) ENGLISH READING PROGRAMS OR
     READING MATERIALS & TECHNIQUES(NOT
     EVALUATION OF SPECIFIC READING TEST
     OR INSTRUCTORS,NOT STUDENT TEST
     SCORES WHEN NOT PART OF EVALUATION
     OF READING PROGRAM,& NOT JUST THE
     CRITERIA OR STANDARDS FOR EVALUAT-
     ION)LIMIT TO PUBS.AVAIL.ERIC/EDRS
D04  FORMAL SCIENCE EDUCATION PROGRAMS
     IN UNIVERSITIES & SECONDARY GRADES
     9-12 IN SOVIET UNION(ALL SCIENCES
     INCL.MATH&ENGIN.?TAPDIF1 FOR MORE
? ?TAPDIF1
D05  VOCATIONAL EDUCATION OF AMERICAN
     INDIAN(HISTORY,DATA,& PROGRAMS TO
     PROVIDE THIS EDUCATION;BUT NOT
     TRAINING MATERIALS TO BE USED)
D06  EVALUATION OF BILINGUAL ELEMENTARY
     ELEMENTARY (K-8) & SECONDARY (9-12)
     SCHOOL PROGRAMS OR TECHNIQUES FOR
     SPANISH & ENGLISH LANGUAGES
D07  FINANCIAL COSTS TO INSTITUTIONS OF
     HIGHER ED. TO IMPLEMENT LEGISLATED
     AFFIRMATIVE ACTION PROGRAMS
D08  EVALUATION OF INDEXING & CATALOGING
     (WORK,METHODS,PRODUCTS,LANGUAGES) TO
     INCLUDE REPRESENTATION & STORAGE OF
     THE INDEX INFO.'INDEXING ' INCLUDES
     ALL FORMS OF TEXT SEARCHING,I.E.
     INDEXING BY TEXT WORDS.'EVALUATION'
     IS TIME (COST, ERROR RATES,RECALL/
     RELEVANCE, AND EVAL.CRITERIA)INDEX-
     ING INCLUDES CITATION INDEXING,
     TITLE WORD INDEXING,OR AUTO INDEX
D09  COLLECTIVE BARGAINING IN LIBRARIES
     OF INSTITUTIONS OF HIGHER EDUCATION
```

Forty-six documents are judged as being relevant to the information-need. In formulation #1, a minimum of online searching effort is expended. The results are as follows:

Formulation #1. (Briefsearch)

Number of citations found = 20

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{15}{46}$ (33% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{15}{20}$ (75% of retrieved documents are relevant)

The Briefsearch performs well in terms of precision, capturing fifteen relevant documents out of the twenty retrieved items. The information seeker, presented with the results of this search does not have to sort through much non-relevant material although 26 relevant items have been missed. On the other hand, formulation #3 which is constructed for high recall retrieves a high percentage of relevant items at the expense of retrieving a high volume of non-relevant items. Recall and precision ratios for all three parts of the search are listed below:

Formulation #3, part I. (Using the general EVALUATION Search Save)

Number of citations found = 250

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{46}{46}$ (100% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{46}{250}$ (18% of retrieved documents are relevant)

Formulation #3, part II. (Using the restricted EVALUATION Search Save)

Number of citations found = 196

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{42}{46}$ (90% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{42}{196}$ (22% of retrieved documents are relevant)

Formulation #3, part III. (Using the abbreviated search for EVALUATION)

Number of citations found = 137

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{39}{46}$ (85% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{39}{137}$ (28% of retrieved documents are relevant)

As formulation #3 proceeds, additional restricting factors are applied to the search. In part I the three facets are made up of two general Search Saves for EVALUATION and PRIMARY EDUCATION, and the occurrences of READING or READER in all subject-conveying fields. Here, all 46 relevant citations are retrieved; however, over 200 retrieved citations are not relevant to the information need.

In part II, the general EVALUATION Search Save and Search Save for PRIMARY EDUCATION are COMBINEd in an AND relationship with a restricted READING set. These limiting factors result in the loss of four relevant items; however, 50 non-relevant items are not included in the final set.

Additional restricting factors applied in part III of formulation #3 result in improved precision in exchange for a lower recall. The third formulation, an example of the successive fractions approach to search strategy, performs well, in this case, for a high recall search in which all relevant citations are desired; but the search objective as expressed by the information seeker is not satisfied since over two-thirds of the output in all three portions of the search contains non-relevant material.

Search formulation #2, made up of restricted sets for EVALUATION and READING, works out surprisingly well in terms of precision and recall values as shown below:

Formulation #2

Number of citations found = 52

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{30}{46} \quad \text{(65\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{30}{52} \quad \text{(58\% of retrieved documents are relevant)}$$

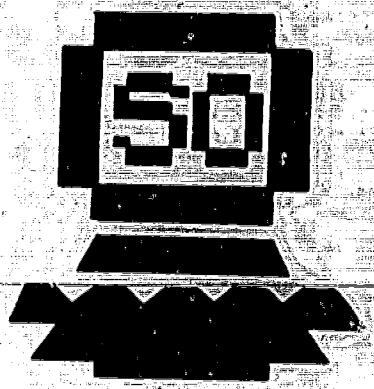
In view of the information-need expressed by the information seeker, the

results of the Briefsearch or formulation #2 may be satisfactory. More relevant documents are retrieved in formulation #2 than the Briefsearch at the expense of drawing out some non-relevant material; however, the Briefsearch serves as a high precision, low cost formulation.

(Reader: Since there are a number of possibilities present in choosing among formulations, which would you select after reading the evaluation?_____)

Relevant citations:

ED110543	ED110511	ED110510	ED110171	ED110170	ED110168	ED109934
ED109686	ED109674	ED109638	ED109632	ED109631	ED109592	ED109150
ED108791	ED108377	ED108201	ED108197	ED108196	ED108192	ED107372
ED107071	ED106750	ED106400	ED105440	ED105012	ED104999	ED104994
ED104925	ED104919	ED104527	ED103008	ED103004	ED100176	ED100175
ED099822	ED099820	ED099812	ED098494	ED098346	ED097645	ED097635
ED097104	ED096781	ED096780	ED096033			



SOCIAL STUDIES

Search Topic #5: Use of school busing to achieve racial integration. (S AN=M08)

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

Two facets comprise the search topic: integration and busing. The concept of school could form another facet; however, since the type of school is not specified in the query, the facet of school will not be applied to the search results unless a large volume of output is retrieved in the formulation below.

Step 4: Compiling the Search Terms

The INTEGRATION aspect can be described by a number of ERIC descriptors, in comparison to BUSING which is only represented by three descriptors, e.g., bus transportation, student transportation, and school buses. Formulations constructed for free text searching can include variant forms of the busing and desegregation since neither term is used singly as a descriptor.

Step 5: Ordering Output

No output specifications are stated in the information need. Thus, the discussion will cover three different search strategies in order to illustrate how each of these approaches is employed to achieve one of three objectives: high recall, high precision, or low cost.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Whether to survey the file or gather additional searching vocabulary, the Briefsearch given below retrieves a high percentage of relevant items at a low cost.

Formulation #1	Comments
1 S STUDENT(W)TRANSPORTATION(C)INTEGRATION	
2 S BUSING(F)DESEGREGATION	
3 C 1 OR 2	After COMBINing the results of the two sets, 23 citations are retrieved. At this point, it is not felt necessary to add the third facet SCHOOL to the formulation and the search is terminated.

As a result of evaluating the Briefsearch, it is decided that the third facet of SCHOOL will not be included in the high recall formulation, an example of the building block approach, which follows:

Formulation #2	Comments
1 S BUS	Using truncation to retrieve the many possible forms of BUS would probably result in a large number of false drops. Each form is typed in separately rather than entering BUS?
2 S BUSES	
3 S BUSING	
4 S BUSED	
5 S BUSSES	
6 S BUSSING	
7 S BUSSED	
8 S STUDENT(C)TRANSPORTATION	Included DESCRIPTORS/Identifiers: BUS TRANSPORTATION, SCHOOL BUSES/Bus Drivers, Bus Driver Training.
9 C 1-8/OR	
10 S RACIAL?	
11 S INTEGRAT?	Included DESCRIPTORS/Identifiers: 22 INTEGRATED & INTEGRATION DESCRIPTORS, SEGREGATED PUBLIC FACILITIES, 7 SEGREGATION DESCRIPTORS, SEGREGATIONIST ORGANIZATIONS/Desegregation Aid, Desegregation Advisory Project
12 S DESEGREGAT?	
13 S SEGREGAT?	
14 C 10-13/OR	
15 C 14 AND 9	

Formulation #3 is an example of the lowest postings facet first approach. Expected to produce high precision, the strategy involves the development and evaluation of the facet BUSING before the application of the second facet INTEGRATION.

Formulation #3	Comments
1 S STUDENT TRANSPORTATION	Using ERIC descriptors to create a set of documents for the facet BUSING should result in high precision since only the assigned index fields are scanned.
2 S BUS TRANSPORTATION	
3 S SCHOOL BUSES	
4 C 1-3/OR	At this point, 63 citations are retrieved. The searcher could end the search now, but decides to add the second facet INTEGRATION which is constructed using free text searching.
5 S INTEGRAT?	
6 S DEGREGAT?	
7 S SEGREGAT?	
8 C 5-7/OR	
9 C 4 AND 8	The BUSING facet is reduced to only those documents containing search terms for INTEGRATION in any of the 7 subject-conveying fields.

Step 8: Evaluating Final Results

The Briefsearch is remarkably accurate in terms of retrieving relevant documents as reflected by high precision. However, all 42 of the relevant items in the file are not retrieved by this formulation. The Briefsearch is favored if high precision and low cost are search objectives.

Briefsearch

Number of citations found = 23

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{21}{42}$$ (50% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{21}{23}$$ (90% of retrieved documents are relevant)

On the other hand, if high recall is desired by the requestor, formulation #2, the building block approach, in which free text searching is used may be preferred. Capturing all the relevant items, the formulation also results in retrieving a high volume of non-relevant citations.

Formulation #2

Number of citations found = 63

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{42}{42}$ (100% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{42}{63}$ (66% of retrieved documents are relevant)

Formulation #3 in which high precision is intended achieves that objective. Developing the BUSING facet by descriptor terms alone is not suitable, in this case, to effect high recall also.

Formulation #3

Number of citations found = 29

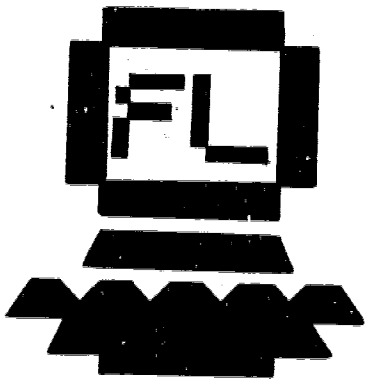
Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{26}{42}$ (60% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{26}{29}$ (90% of retrieved documents are relevant)

There is a "tradeoff" in each of the formulations. For example, formulation #3 achieves high recall at the risk of retrieving much non-relevant material. At the expense of recall, the Briefsearch is a low cost, high precision formulation. Compare your search formulations with those given. Check to see whether you have compromised in order to achieve high recall, or vice versa.

Relevant citations;

EJ 121907	EJ 121906	EJ 121889	EJ 121888	EJ 121886	EJ 121885
EJ 120204	EJ 115861	EJ 115858	EJ 115853	EJ 115849	EJ 115809
EJ 115252	EJ 114427	EJ 113526	EJ 113504	EJ 111820	EJ 110130
EJ 110129	EJ 110036	EJ 108492	EJ 108055	EJ 107317	EJ 107301
EJ 106768	EJ 105109	EJ 103409	EJ 103405	EJ 103347	ED 110551
ED 110533	ED 109745	ED 106425	ED 106397	ED 105554	ED 103546
ED 103288	ED 102279	ED 102242	ED 101045	ED 101018	ED 099258



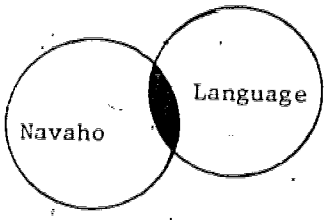
LANGUAGES AND LINGUISTICS

Search topic #6: Navaho language textbooks or grammars (material in Navaho or useful for teaching Navaho, or about Navaho linguistics). (S AN=S05)

Search Objective: The information seeker is interested in gathering as much information as possible on this topic at the lowest cost.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

Two concepts are present in the topic: Navaho and language. The diagram below illustrates the search logic; the shaded section represents the portion of the data base intended for retrieval.



Step 4: Compiling the Search Terms

There is sufficient terminology for expressing the facet LANGUAGE in assigned index terms. The NAVAHO facet can be represented by a single descriptor, so that related terms such as American Indian Languages or American Indian Culture could be introduced to the formulation; however, these terms may retrieve items treating American Indian Languages as well as the Navaho language.

Step 5: Ordering Output

Selecting an approach to search strategy should take into consideration the search objectives: high recall and low cost. The most specific facet first approach, since it permits the searcher to terminate the search before introducing all facets to the formulation, is preferred as well as the Briefsearch. Developing

a high precision formulation is also necessary as the cost of the search involves offline prints as well as online connect time.

Step 6: Conceptualizing the Search as Input to the Retrieval System

A Briefsearch is constructed by taking variant spellings of Navaho and combining them with the search term Language using the AND operator. The (C) limiter could be used instead of the AND operator; however, it cannot be used in conjunction with truncation.

Briefsearch	Comments
1 S NAVA?O? ?	Different spellings of NAVAHO are easily retrieved employing DIALOG's truncation feature; since the root of the term is not common to many words, few false drops should occur.
2 S LANGUAGE?	Since 49 citations are retrieved after developing the most specific facet, NAVAHO, the searcher decides to apply the second facet rather than risk retrieving much nonrelevant material.
3 C 1 AND 2	

The Briefsearch retrieves 20 citations; the searcher could end the search at this point or continue developing these statements into a more complex formulation. Formulation #2 incorporates terms taken from the search topic into the LANGUAGE facet. A greater volume of output can be expected as citations are retrieved which contain not only Language, but Textbooks or Books, or Grammars, etc.

Formulation #2	Comments
4 S BOOK?	Formulation #2 uses the groundwork provided by the Briefsearch; in this way, the searcher does not have to backtrack to create sets for NAVAHO and LANGUAGE.
5 S GRAMMAR?	
6 S LINGUISTIC?	
7 S TEXTBOOK?	
8 C 2 OR 4 OR 5 OR 6 OR 7	
9 C 1 AND 8	

(Reader: After executing your own search and reviewing the sample formulations, which formulation would you select? Keep in mind that the requestor is concerned about cost. _____)

Step 8: Evaluating Final Results

The Briefsearch manages to capture eight of the nine relevant items in the ONTAP file. Only three search statements are entered in this formulation keeping online connect time at a minimum.

Briefsearch

Number of citations found = 20

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{8}{9} \quad \text{(89\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{8}{20} \quad \text{(40\% of retrieved documents are relevant)}$$

Just as the Briefsearch retrieves only eight of the nine relevant documents, the second formulation fared the same; however, three additional non-relevant citations are picked up as a result of including more search terms into the development of the LANGUAGE facet.

Formulation #2

Number of citations found = 23

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{8}{9} \quad \text{(89\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{8}{23} \quad \text{(35\% of retrieved documents are relevant)}$$

Both formulations fail to retrieve the same item (ED 104168) judged as being relevant to the topic; curious to find out why the document is missed, we printed it in full format. No portion of the record mentions specifically the Navaho language; it deals generally with North American Indian languages, so that its relevance to the topic may be questioned.

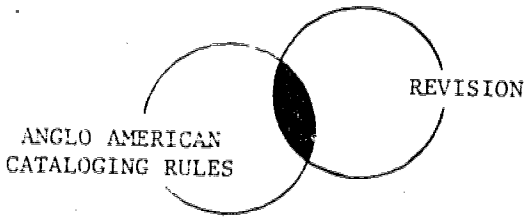
Relevant citations:

EJ 111077 EJ 111075 ED 108801 ED 104168 ED 102820 ED 102585
ED 101549 ED 101152 ED 100136

Search topic #7: Revision of the Anglo-American Cataloging Rules. (S AN=SO4)

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

Two concepts are present in the topic: REVISION and ANGLO-AMERICAN CATALOGING RULES. The shaded section in the diagram below indicates the portion of the data base intended for retrieval.



Step 4: Compiling Search Terms

Anglo American Cataloging Rules, AACR or revision are not included as descriptors in the ERIC Thesaurus so much of the formulation will consist of free text searching. A high precision search would be restricted to the identifier and title fields.

Step 5: Ordering Output

No output specifications are identified in the query. The formulations give examples of the Briefsearch, building block approach, and a high recall formulation in which the identifier and title fields are searched.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Formulation #1 serves as an example of the building block approach. All subject-conveying portions of the ERIC record are searched for occurrences of

the search terms as free text searching is employed.

Formulation #1	Comments
1 S AACR	
2 S ANGLO(1W)CATALOGING	Since the Rules can be spelled with or without a hyphen, this statement retrieves both forms
3 S REVIS?	
4 C (1 OR 2) AND 3	

A high precision search uses nearly the same terms but scans the identifier field for the ANGLO-AMERICAN facet.

Formulation #2	Comments
1 S AACR/ID	
2 S ANGLO AMERICAN CATALOGING RULES	After EXPANDing ANGLO AMERICAN as an identifier, SELECT from the alphabetical listing instead of typing it in yourself.
3 S REVISION	
4 C (1 OR 2) AND 3	

A Briefsearch might assume the following form:

Formulation #3: Briefsearch	Comments
1 S ANGLO(1W)CATALOGING(C)REVISION	The (C) limiter saves another step for SELECTing and then COMBINing the two results.

Step 8: Evaluating Final Results

Formulation #1

Number of citations found = 5

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{3}{3}$$ (100% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{3}{5}$$ (60% of retrieved documents are relevant)

Formulations #2 and #3

Number of citations found = 3

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{3}{3} \quad \text{(100\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{3}{3} \quad \text{(100\% of retrieved documents relevant)}$$

After reviewing the recall and precision scores summarized above, it is clear that no one formulation is favored; all perform well and are efficient and accurate.

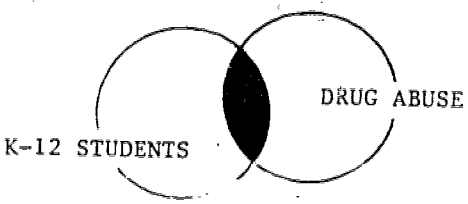
Relevant citations:

EJ 121112 ED 108684 ED 100292

Search topic #8: Drug abuse including alcohol, among students of elementary and secondary schools, grades K-12. (S AN=M07)

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

The search topic can be broken down into two facets: Drug abuse and K-12 students. The diagram below shows that the intersection of the two facets is the portion of the data base intended for retrieval.



Step 4: Compiling the Search Terms

This search is simplified by the availability of ERIC descriptors for DRUG ABUSE and for student grade level. A high recall search may contain the names of specific drugs such as marihuana or LSD.

Step 5: Ordering Output

It is anticipated that the volume of literature on the topic is quite substantial; consequently, both facets will probably be introduced to the formulation. The building block approach is used in all three formulations below for a high recall and high precision search, and Briefsearch; the building block approach is characterized by the development and union of all search facets.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Each search statement in the Briefsearch given below is an example of the

building block approach as both facets, DRUG ABUSE and K-12 STUDENTS are present. Search terms chosen for the Briefsearch are derived from ERIC descriptors having high postings or used frequently in descriptor phrases.

Briefsearch	Comments
1 S ALCOHOL(C)SCHOOL(C)STUDENTS	The (C) limiter is used rather than the AND in order to reduce the number of individual SELECT and COMBINE operations
2 S DRUG(W)ABUSE(C)HIGH(W)SCHOOL	
3 S DRUG(W)ABUSE(C)ELEMENTARY	
4 C 1-3/OR	

A high precision formulation relying solely on the descriptor phrase Drug Abuse as a descriptor, identifier, or title word, to create the DRUG ABUSE facet and descriptors for grade levels is shown below:

Formulation #2

1 S DRUG(W)ABUSE/DE,ID,TI

2 S ELEMENTARY SCHOOL STUDENTS

3 S HIGH SCHOOL STUDENTS

4 S SECONDARY SCHOOL STUDENTS

5 C (2 OR 3 OR 4) AND 1

Comprehensive Search Saves for grade levels K-8 (Search Save #2, see appendix) and grade levels 9-12 (Search Save #3, see appendix) are used in the following high recall formulation to create the K-12 STUDENTS facet. The DRUG ABUSE facet is comprised of individual drug names in addition to general terms such as narcotics, drug addiction, etc.

Formulation #3		Comments
1 EXECUTE_____	(Search Save for ELEMENTARY EDUCATION)	Statements 1 and 2 create the set for K-12 STUDENTS
2 EXECUTE_____	(Search Save for SECONDARY EDUCATION)	

Formulation #3 (cont.)

3 S ALCOHOL?
4 S DRUG(W)ABUSE
5 S DRUG(W)ADDICTION
6 S LSD
7 S LYSERGIC
8 S MARI?UANA?
9 S NARCOTIC?
10 C 3-9/OR
11 C 10 AND (1 OR 2)

Comments

Included DESCRIPTORS and Identifiers: ALCOHOL EDUCATION, ALCOHOLIC BEVERAGES, ALCOHOLISM, Alcohol, 4 multi-term alcohol identifiers, DRUG ABUSE, 5 multi-term drug abuse identifiers, DRUG ADDICTION, LYSERGIC ACID DIETHYLAMIDE, MARIHUANA, NARCOTICS, 2 multi-term Narcotics identifiers

DIALOG's truncation feature allows variant spellings of the term to be retrieved

Step 8: Evaluating Final Results

Both the Briefsearch and formulation #2 are favored in instances when high precision is required. Simple, efficient and requiring little typing, the searches retrieve few non-relevant documents.

Briefsearch

Number of citations found = 44

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{39}{65}$ (60% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{39}{44}$ (89% of retrieved documents are relevant)

Formulation #2

Number of citations found = 11

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{11}{65}$ (17% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{11}{11}$ (100% of retrieved documents are relevant)

Search formulation #3 is preferable for high recall searches. More online connect time is necessary than formulations #1 or #2, as more terms are entered and Search Saves are processed. However, all relevant documents are retrieved

with a minimum of nonrelevant items.

Formulation #3

Number of citations found = 75

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} = \frac{65}{65}$ (100% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}} = \frac{65}{75}$ (87% of retrieved documents are relevant)

Relevant citations:

EJ119196	EJ118984	EJ118982	EJ118979	EJ117327	EJ116766	EJ115585
EJ115584	EJ115581	EJ115577	EJ114039	EJ114035	EJ114034	EJ114033
EJ114032	EJ114029	EJ110527	EJ110132	EJ108320	EJ107353	EJ107314
EJ105590	EJ105584	EJ105583	ED110574	ED109914	ED109844	ED109165
ED109069	ED109063	ED108016	ED107899	ED107664	ED107663	ED107618
ED107572	ED107550	ED106987	ED106719	ED106705	ED104786	ED104620
ED104006	ED103756	ED103746	ED103461	ED103361	ED102658	ED101866
ED101499	ED101464	ED101256	ED100642	ED100610	ED100597	ED099720
ED099398	ED099354	ED099338	ED098465	ED098181	ED097604	ED097234
ED096577	ED095691					

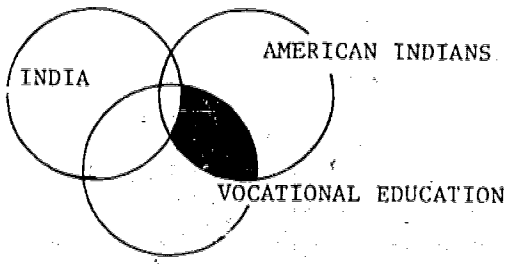
Search topic #9: Vocational education of the American Indian (history, data and programs to provide this education but not training materials to be used in these educational programs. (S AN=D05)

Search Objective: The requestor wants all the relevant material in the ONTAP file.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

Two concepts are present in the topic: Vocational Education and American Indians. Although the requestor is not interested in training materials, this aspect of the search will be difficult to represent in the formulation without eliminating some relevant items which contain information on industrial training. Errors in indexing the concept AMERICAN INDIANS may occur so that a set representing INDIA will be created and joined in a NOT relationship with the facet for AMERICAN INDIANS. In this way, material dealing with American Indians but indexed as INDIANS would be retrieved. (INDIANS is used as an ERIC descriptor to describe "natives of India or of the East Indies," according to the scope note in the ERIC Thesaurus).

The search logic is illustrated in the diagram below; the shaded area is the portion of the data base intended for retrieval. Note that the intersection of all three sets is not wanted, rather the intersection of the two sets for VOCATIONAL EDUCATION and AMERICAN INDIANS is desired in all cases where the term INDIA does not occur.



Step 4: Compiling the Search Terms

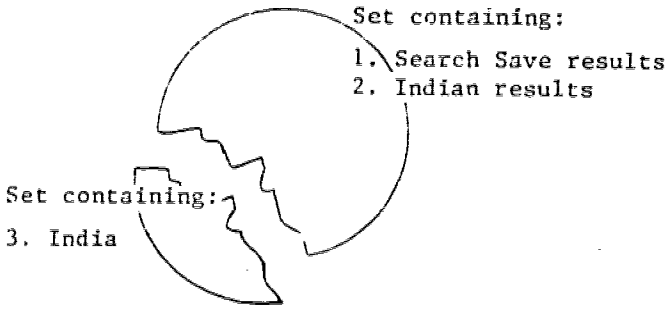
The VOCATIONAL EDUCATION aspect is well represented in the ERIC Thesaurus by a host of terms. Use can be made of the online thesaurus instead of typing in a number of terms and risking typing errors. For a high recall search, several expressions must be entered to represent AMERICAN INDIANS, particularly individual tribal names. Formulation #2 below employs Search Save #5 (see appendix) on general Indian names; however, a more comprehensive Search Save containing nearly all tribal names of North American Indians could be devised for high recall searches.

Step 5: Ordering Output

The search objective, to retrieve all relevant citations in ONTAP, plays a part in determining what approach to search strategy should be employed. Since neither facet, American Indians nor Vocational Education, could represent the topic alone without containing much non-relevant material, it is expected that both facets will be involved in the formulation. Thus, the building block approach appears to be a suitable choice; however, the successive fractions approach could also be employed to execute a comprehensive subsearch for American Indians; in this way, items indexed erroneously and/or dealing with American Indians but not referring to them specifically are retrieved.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Formulation #1 is an example of the successive fractions approach. In the first four search statements, the facet AMERICAN INDIANS is built by using Search Save #5 and the search terms India and Indian; then, the set for INDIA is partitioned from the AMERICAN INDIAN set. The following diagram illustrates this process; the shaded area represents that portion of the data base which is treated in subsequent statements of the formulation; the unshaded area is, literally, discarded.



The full formulation is given below:

Formulation #1		Comments
1	.EXECUTE_____ (Search Save for AMERICAN INDIANS)	
2	S INDIAN	Statements 2-4 draw out material on INDIAN and INDIA. The set for INDIA is then COMBINEd in a NOT relationship with the AMERICAN INDIAN facet in order to eliminate items on India which may be non-relevant
3	S INDIA	
4	C (1 OR 2) NOT 3	
5	S INDUSTRIAL(W)ARTS	Statements 5-9 comprise the facet of VOCATIONAL EDUCATION
6	S INDUSTRIAL(W)EDUCATION	
7	S VOCATION?	
8	S INDUSTRIAL(W)TRAINING	
9	C 5-8/OR	
10	C 4 AND 9	

The Briefsearch formulation, a neat and efficient single search statement, retrieves all but one relevant item.

Briefsearch		Comments
1	S AMERICAN(F)INDIANS(C)VOCATIONAL	DIALOG processes word proximity statements which range in length from a few characters to as many as fit on a single typed line

Step 8: Evaluating Final Results

Before reviewing the evaluation of the Briefsearch and the high recall formulation, how well does your search perform in terms of recall and precision? Twenty-one citations are judged as being relevant. Did you retrieve all the relevant documents?

The results of the high recall formulation reveal that all relevant citations are retrieved.

Formulation #1

Number of citations found = 35

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{21}{21}$$ (100% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{21}{35}$$ (60% of retrieved documents are relevant)

On the other hand, the Briefsearch failed to recall one relevant citation. An error in indexing accounts for this, as the descriptor Indians which refers to the "Natives of India or of the East Indies" according to its scope note, is assigned to citation ED 097483 dealing with Indians in Arizona. In order to retrieve this item, it is necessary to execute the NOT operation as described in formulation #1 or enter individual tribal names of North American Indians. Recall and precision for the Briefsearch are given below:

Briefsearch

Number of citations found = 25

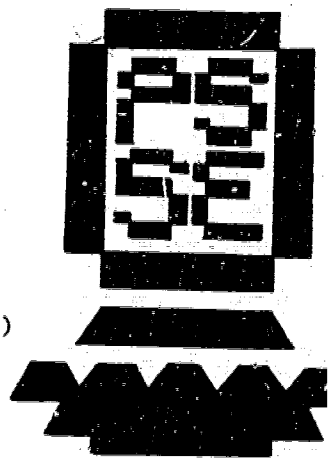
$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{20}{21}$$ (95% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{20}{25}$$ (80% of retrieved documents are relevant)

The question of which formulation is preferred in this case is left to you, the searcher.

Relevant citations:

EJ 113861	EJ 112180	ED 109445	ED 108822	ED 108816	ED 108797	ED 108464
ED 107424	ED 107396	ED 107395	ED 101313	ED 101167	ED 100542	ED 100541
ED 100539	ED 100437	ED 098000	ED 097483	ED 097163	ED 096073	ED 095410



Search topic #10: Jean Piaget's theories, and thought processes or language development of children. (S AN=M03)

Search Objective: All relevant citations are wanted by the requestor; no journal articles are desired.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

Piaget represents one facet of the search. Piagetian theories, language development and thought processes comprise aspects of the second facet (this second facet is referred to as THEORIES in this discussion). Children forms the third facet as the requestor desires information on theories of Children, thought processes of Children, or language development of Children.

Step 4: Compiling the Search Terms

To incorporate Piagetian theories in the formulation, the searcher must be aware of the vocabulary and terminology associated with Piaget's work. The need to include many variant or synonymous expressions for the theories adds to the complexity of this search. The ERIC Thesaurus and Thesaurus of Psychological Terms are valuable sources for gathering such search terms. In search formulation #4, terminology for thought processes, language development and Piagetian theories make up a substantial portion of the search.

Step 5: Ordering Output

Since the output is limited to publications available from EDRS, the second step in every formulation is the LIMIT command; this restricts the results of the first facet to the appropriate body of documents and allows the searcher to monitor preliminary results in the light of one of the search objectives.

Beginning with the most specific aspect of the search, PIAGET, the searcher may decide to terminate the search without adding another facet to the formulation. Examples of the most specific aspect first, building block and Briefsearch are given. The COMMENTS section of each formulation explains the searcher's decisions during the preliminary evaluation process.

Step 6: Conceptualizing the Search as Input to the Retrieval System

A Briefsearch, performed initially before a complex formulation, gives a rough estimation of the size of the final output. Records from Briefsearch results can be reviewed in order to collect additional searching vocabulary.

Briefsearch

1 S PIAGET(C)COGNITIVE
2 S PIAGET(C)CHILDREN
3 C 1 OR 2
4 LIMIT 3/AVAIL

At this point 72 citations are retrieved. Searching on the most specific facet, PIAGET, and comparing the results of this statement with that of the Briefsearch gives us a general idea how large the final set may be.

Formulation #2	Comments
1 S PIAGET?	Included DESCRIPTORS and identifiers: Piaget, 17 multi-term Piaget identifiers, Piagetian, 9 multi-term Piagetian identifiers.
2 LIMIT 1/AVAIL	This search retrieves 93 citations. Since the requestor wants all relevant items, the searcher may consider terminating the search at this point, having determined that the volume of output is small enough for manual review.

The addition of the facet CHILDREN to the formulation reduces the output to 77 citations. This size of output may be considered sufficient to forego the introduction of another facet to the formulation.

Formulation #3	Comments
3 . EXECUTE _____ (Search Save for CHILDREN)	Search Save #4 (see appendix) is used for the CHILDREN facet
4 C 2 AND 3	

The fourth formulation is comprised of two facets, THEORIES and PIAGET. It requires much typing so that using the online thesaurus may be favored instead of employing free text searching.

Formulation #4	Comments
1 S PIAGET?	Just like in formulation #3. the first two statements are taken from formulation #2 and make up the entire facet PIAGET.
2 LIMIT 1/AVAIL	
3 S AFFECTIVE(W)BEHAVIOR	Included DESCRIPTORS/Identifiers: AFFECTIVE BEHAVIOR
4 S COGNITIVE(W)ABILITY	COGNITIVE ABILITY
5 S COGNITIVE(W)ABILITIES	Cognitive Abilities Test
6 S COGNITIVE(W)DEVELOPMENT	COGNITIVE DEVELOPMENT/Cognitive Development Research Tools
7 S COGNITIVE(W)DEVELOPMENTS	
8 S COGNITIVE(W)OBJECTIVE	
9 S COGNITIVE(W)OBJECTIVES	COGNITIVE OBJECTIVES
10 S INTELLECTUAL(W)DEVELOPMENT	INTELLECTUAL DEVELOPMENT
11 S INTELLECTUAL(W)DEVELOPMENTS	
12 S MENTAL(W)DEVELOPMENT	MENTAL DEVELOPMENT
13 S PERCEPTUAL(W)DEVELOPMENT	PERCEPTUAL DEVELOPMENT
14 S ECHOLALIA	ECHOLALIA
15 S FLES	FLES, FLES GUIDES, FLES MATERIALS, FLES PROGRAMS, FLES TEACHERS/FLES Objectives, FLES Program
16 S READING(W)HABIT	
17 S READING(W)HABITS	READING HABITS
18 S HUMAN(W)DEVELOPMENT	HUMAN DEVELOPMENT/3 multi-term Human Development Identifiers
19 S LANGUAGE?	LANGUAGE, over 60 multi-term LANGUAGE DESCRIPTORS, 31 LANGUAGES DESCRIPTORS/ 50 multi-term Language Identifiers
20 S NUCLEATION	NUCLEATION (LANGUAGE LEARNING)
21 S PSYCHOLINGUISTIC?	PSYCHOLINGUISTICS
22 S READING(W)DEVELOPMENT	READING DEVELOPMENT
23 S LINGUISTIC?	LINGUISTICS, 13 multi-term LINGUISTICS and LINGUISTIC DESCRIPTORS/26 multi-term Linguistic and Linguistics Identifiers

Formulation #4 (Cont.)		Comments
24	S SPEECH(W)DEVELOPMENT	RETARDED SPEECH DEVELOPMENT
25	S SPEECH(W)HABITS	SPEECH HABITS
26	S VERBAL(W)DEVELOPMENT	VERBAL DEVELOPMENT
27	S VOCABULARY(W)DEVELOPMENT	VOCABULARY DEVELOPMENT
28	S HANDWRITING(W)DEVELOPMENT	HANDWRITING DEVELOPMENT
29	C 3-28/OR	Cumulative set for cognitive and language development
30	C 29 AND 2	

Imposing the facet PIAGETIAN THEORIES reduces the final set to 71 citations. At this point, the searcher terminates the search in order to avoid losing relevant material. If a decision is made to include the CHILDREN facet, use could be made of the Search Save demonstrated in formulation #3.

30 .EXECUTE _____ (Search Save for CHILDREN)

31 C 30 AND 29

After the CHILDREN facet is incorporated, the total number of retrieved citations is 60.

(Reader: In light of the requestor's desire to retrieve all relevant citations regardless of cost, which strategy do you prefer? _____)

Step 8: Evaluating Final Results

Formulation #2 in which the most specific facet of the topic is searched retrieves all 81 relevant citations. It is the only formulation of those given which achieves 100% recall.

Formulation #2

Number of citations found = 93

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{81}{81}$$ (100% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{81}{93}$$ (87% of retrieved documents are relevant

The Briefsearch achieves high precision and high recall but does not retrieve all items relevant to the topic. But, used as a brief survey of the file, it aids in determining whether to impose a second facet to the formulation.

Briefsearch

Number of citations found = 72

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{66}{81}$ (81% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{66}{72}$ (92% of retrieved documents are relevant)

Two facets, CHILDREN and PIAGET, of the topic make up formulation #3. Imposing the CHILDREN aspect into the formulation eliminates 14 relevant documents from the output.

Formulation #3

Number of citations found = 77

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{67}{81}$ (83% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{67}{77}$ (87% of retrieved documents are relevant)

All facets of the topic are included in the building block approach, formulation #4. If the search is terminated before the CHILDREN facet is included in the formulation, 71 citations are retrieved; after the imposition of the CHILDREN facet, 60 citations are retrieved.

Formulation #4. ("Before" CHILDREN facet, statements 1-31)

Number of citations found = 71

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{71}{81}$ (88% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{71}{71}$ (100% of retrieved documents are relevant)

Formulation #4. ("After" CHILDREN facet, statements 1-33)

Number of citations found = 60

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{60}{81}$$ (75% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{60}{60}$$ (100% of retrieved documents are relevant)

As more facets are added to the formulation in the third and fourth examples, the recall ratio decreases. Since high recall is one of the search objectives, formulation #2 is preferred. Not only does it achieve a high recall score, but it is efficient and rapid in terms of online connect time and pre-search preparation.

Relevant citations:

ED 110469	ED 110190	ED 110186	ED 110173	ED 110159	ED 109196	ED 109149
ED 108957	ED 108893	ED 108864	ED 108862	ED 108751	ED 108748	ED 108745
ED 108404	ED 107562	ED 107385	ED 107366	ED 106721	ED 106147	ED 106131
ED 106122	ED 106117	ED 106042	ED 105979	ED 105976	ED 104739	ED 104552
ED 104530	ED 104516	ED 104051	ED 103296	ED 103133	ED 103129	ED 103127
ED 103106	ED 103100	ED 103095	ED 102240	ED 102111	ED 101849	ED 101847
ED 101837	ED 101565	ED 101488	ED 100704	ED 100536	ED 100529	ED 100528
ED 100521	ED 100510	ED 100488	ED 100485	ED 099811	ED 099767	ED 099415
ED 099196	ED 099122	ED 099112	ED 099098	ED 098793	ED 098517	ED 098490
ED 098082	ED 097991	ED 097984	ED 097977	ED 097970	ED 097959	ED 097957
ED 097379	ED 097261	ED 097129	ED 097110	ED 097109	ED 097102	ED 096016
ED 096000	ED 095999	ED 095997	ED 095996			

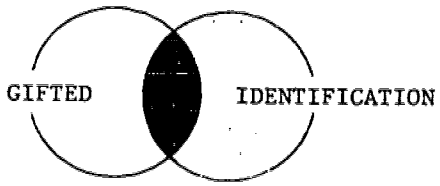
HANDICAPPED AND GIFTED CHILDREN

Search topic #11: Identification of the Gifted.

Search objective: The requestor needs immediate results.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

The search topic is broken down into two facets: identification and gifted. The diagram below illustrates search logic:



Step 4: Compiling the Search Terms

Search terms can be gathered from the ERIC Thesaurus, Roget's Thesaurus, and the Thesaurus of Psychological Terms. However, the requestor's desire for immediate results curtails lengthy pre-search preparation on the part of the searcher. The search topic is comprised of two heavily posted descriptors; these terms also represent both search facets. Thus, the searcher can use the terminology of the search topic as search terms rather than spend time examining various thesauri.

Step 5: Ordering Output

The searcher, obliged to produce immediate results, must curtail pre-search preparation; selecting the citation pearl growing approach to search strategy allows the searcher to review preliminary results of the search in order to gather additional search terms so that preparation before the online interaction

is not essential. Moreover, the searcher can revise search strategy when a search term retrieves non-relevant material.

High precision is an unstated objective in this search as the requestor's demand for immediate results requires online printing of the output. Online printing of a large percentage of non-relevant items is costly in terms of online connect time and requires that the information seeker spend time wading through non-relevant material rather than concentrating on the relevant items. In addition, online printing of the format that provides only the information essential to the requestor is suggested.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Using the citation pearl growing approach, the searcher begins with a very direct search on the most specific search terms. Expressing the topic in the terms used by the requestor, identification and gifted, is limited to retrieving citations whose titles contain these terms. The searcher then prints a few citations in format #2, checking the descriptor field in order to gather other search terms.

Formulation #1	Comments
1 S GIFTED(F)IDENTIFICATION/TI	Five citations are retrieved; the searcher, confident that all five are relevant, checks the index terms in order to add more search terms to the formulation.

TYPE 1/2/1-2

EJ103421 UD503083
THE IDENTIFICATION OF MENTALLY GIFTED, "DISADVANTAGED" STUDENTS AT THE EIGHTH GRADE LEVEL
FITZ-GIBBON, CAROL T.
JOURNAL OF NEGRO EDUCATION, 43, 1, 53-66 W 74
DESCRIPTORS: *JUNIOR HIGH SCHOOL STUDENTS/ *TALENT IDENTIFICATION/ *DISADVANTAGED YOUTH/ *URBAN SCHOOLS/ NEGRO STUDENTS/ GIFTED/ TALENTED STUDENTS/ ECONOMICALLY DISADVANTAGED/ PREDICTIVE MEASUREMENT
IDENTIFIERS: *CALIFORNIA

ED108278* EA007191
GIFTED STUDENTS: IDENTIFICATION TECHNIQUES AND PROGRAM ORGANIZATION.
DOOB, HEATHER S.
EDUCATIONAL RESEARCH SERVICE, WASHINGTON, D.C.
APR 75 61P.
AVAILABLE FROM: EDUCATIONAL RESEARCH SERVICE, INC., 1815 NORTH FORT MYER DRIVE, ARLINGTON, VIRGINIA 22209 ($5.00, PAYMENT MUST ACCOMPANY ORDERS $10.00 OR LESS)
DOCUMENT NOT AVAILABLE FROM EDRS
DESCRIPTORS: ABILITY IDENTIFICATION/ BIBLIOGRAPHIES/ ELEMENTARY SECONDARY EDUCATION/ *GIFTED/ *PROGRAM CONTENT/ *PROGRAM DESCRIPTIONS/ *SUPERIOR STUDENTS/ *TALENTED STUDENTS/ TALENT IDENTIFICATION

Gathering the term "Talent" from the descriptor field, the searcher revises the strategy in the second formulation restricting all search terms to occurrences in the descriptor, identifier-or-title fields.

Formulation #2

1 S GIFTED/DE,ID,TI

2 S TALENT/DE,ID,TI

3 S TALENTED/DE,ID,TI

4 S IDENTIFICATION/DE,ID,TI

5 C (1 OR 2 OR 3) AND 4

Employing the same search terms as formulation #2, the following high recall formulation searches all subject-conveying fields rather than being limited to the title and assigned index term fields.

Formulation #3

1 S GIFTED?

2 S TALENT?

3 S IDENTIF?

4 C (1 OR 2) AND 3

Comments

Included DESCRIPTORS: GIFTED, IDENTIFICATION, ABILITY IDENTIFICATION, 4 multi-term IDENTIFICATION DESCRIPTORS, TALENT IDENTIFICATION, TALENT, 3 multi-term TALENT DESCRIPTORS, TALENTED STUDENTS

The Briefsearch, a single search statement, resembles the wording of the search topic.

Briefsearch

1 S IDENTIFICATION(F)GIFTED

(Reader: Keeping in mind that high precision is a search objective as well as delivering output to the requestor as soon as possible, which formulation above fulfills both objectives? _____)

Step 8: Evaluating Final Results

The answer set for the search topic is not available in the ONTAP file. The results of your search can be compared with the list of ERIC documents (cited by accession number) concluding this discussion.

The performance of the first formulation is not evaluated. Intended as a demonstration of how the citation pearl growing approach can be employed to gather search terms and relevant items, the formulation is actually integrated into the second example.

Formulation #2 achieves high recall, as intended, but fails to retrieve all relevant items.

Formulation #2

Number of citations found = 18

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{16}{26}$ (61% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{16}{18}$ (89% of retrieved documents are relevant)

Formulation #3, an example of a high recall formulation, retrieves all relevant items; however, the output contains many non-relevant items, so that this formulation is not the preferred mode for representing the topic.

Formulation #3

Number of citations found = 45

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{26}{26}$ (100% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{26}{45}$ (58% of retrieved documents are relevant)

The evaluation of the Briefsearch follows. Although it retrieves a larger number of relevant items than the second formulation, eight non-relevant citations

are included in the output so that formulation #2 is favored for representing the topic in light of the search objectives.

Briefsearch

Number of citations found = 30

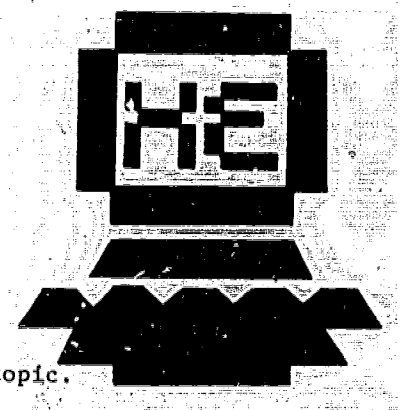
Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{22}{26}$ (85% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{22}{30}$ (73% of retrieved documents are relevant)

Relevant citations:

EJ 110029	EJ 109349	EJ 104281	EJ 104226	EJ 103421	ED 109867
ED 109863	ED 108278	ED 107746	ED 106140	ED 105687	ED 105329
ED 104971	ED 104965	ED 104095	ED 104094	ED 104039	ED 102795
ED 102773	ED 102772	ED 102771	ED 102765	ED 100102	ED 100072
ED 097788	ED 096784				

HIGHER EDUCATION

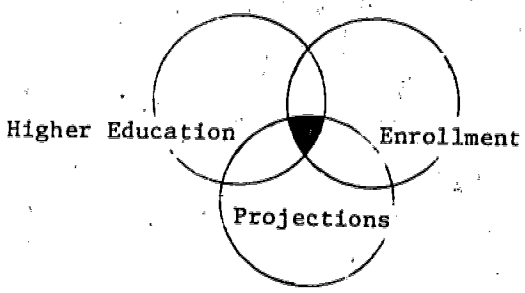


Search topic #12: Higher education enrollment projections for the near future (1975-1995).

Search Objective: To retrieve all journal literature on this topic.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

There are three concepts present in the search topic: higher education, enrollment and projections. Search logic is illustrated below:



Step 4: Compiling the Search Terms

The ERIC Thesaurus is suitable for compiling search terms for the facets HIGHER EDUCATION and ENROLLMENT. Terms synonymous with PROJECTIONS can be gathered from Roget's Thesaurus as well as a dictionary since the descriptor, Project, is associated with Audio-Visual equipment.

Step 5: Ordering Output

The requestor only desires journal literature so the searcher's first task during the online interaction is to apply the LIMITALL feature. Thus, confronted with decisions when online, the searcher can assess preliminary results more accurately. If the LIMIT capability is appended to the formulation as a final step, the size or contents of the final set may not be satisfactory, causing the searcher to "backtrack" or reformulate the search.

High recall and Briefsearch formulations are covered in Step 6. The building block approach is the favored mode for representing search strategy as all three facets have heavily posted search terms. Restricting the formulation to two facets results in retrieving a high volume of non-relevant material.

Step 6: Conceptualizing the Search as Input to the Retrieval System

In order to compile additional search terms or to perform a survey of the file, a Briefsearch can be formulated.

Briefsearch	Comments
LIMITALL/EJ	282 citations are retrieved as a result of the first search statement making it necessary to impose the third facet into the formulation.
1 S HIGHER(W)EDUCATION(F)ENROLLMENT	
2 S PROJECTION?	
3 C 1 AND 2	

The high recall formulation employs Search Save #10(see appendix) to represent the HIGHER EDUCATION facet. As an example of the building block approach, the following formulation is not terminated until all facets are introduced into the strategy.

Formulation #2	Comments
LIMITALL/EJ	
1 S .EXECUTE #_____ (Search Save for HIGHER EDUCATION)	
2 S ENROLLMENT?	The decision to impose the third facet is made after COMBINing statements 1 AND 2, and retrieving over 500 citations.
3 S PREDICT?	
4 S PROJECTION?	
5 S FORECAST?	
6 S FUTUR?	
7 C 3-6/OR	
8 C 1 AND 2 AND 7	

Step 8: Evaluating Final Results

The answer set to this topic is not stored in the ERIC ONTAP file; however,

you can compare your search results with the list of relevant citations concluding this discussion. A citation is determined relevant only if it makes enrollment projections, and not if it treats enrollment projections from a historical perspective or is about making enrollment projections.

Formulation #2 achieves 100% recall, retrieving all 22 relevant documents in the ERIC ONTAP file.

Formulation #2

Number of citations found = 27

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{22}{22} \quad \text{(100\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{22}{27} \quad \text{(81\% of retrieved documents are relevant)}$$

The Briefsearch retrieves 17 items, all of which are relevant. It would be preferred in cases when high precision and/or low cost are search objectives.

Briefsearch

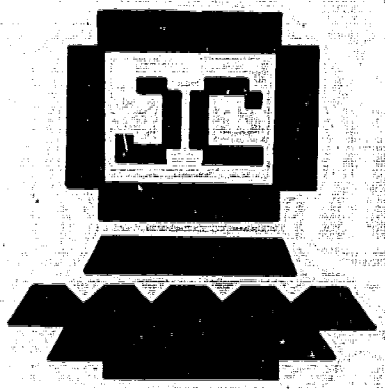
Number of citations found = 17

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{17}{22} \quad \text{(77\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{17}{17} \quad \text{(100\% of retrieved documents are relevant)}$$

Relevant citations:

EJ 117830	EJ 114796	EJ 114396	EJ 114216	EJ 114215	EJ 114214
EJ 112284	EJ 109179	EJ 108126	EJ 108036	EJ 107561	EJ 107550
EJ 107306	EJ 106061	EJ 105868	EJ 104496	EJ 104491	EJ 104490
EJ 104441	EJ 102836	EJ 102829	EJ 102812		



JUNIOR COLLEGES

Search topic #13: Junior and community college enrollment projections for the near future (1975-1995).

Search Objective: The information seeker wants the output restricted to ERIC report literature so that all journal articles are excluded.

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

This topic, like the one preceding on higher education enrollment projections, is comprised of three facets: Junior Colleges, Enrollment and Projections.

Step 4: Compiling the Search Terms

The ERIC Thesaurus, Roget's Thesaurus, and a dictionary are suitable sources for gathering search terms. Instead of formulating a subsearch for the JUNIOR COLLEGES facet, the ERIC Clearinghouse on Junior Colleges could be substituted, as that clearinghouse processes material dealing specifically with Junior Colleges.

Step 5: Ordering Output

Only ERIC report literature is desired so the LIMITALL command is applied as one of the first activities during the online interaction. The building block and successive fractions approaches discussed below, are examples of high recall and high precision formulations, respectively.

Step 6: Conceptualizing the Search as Input to the Retrieval System

In the successive fractions approach, the LIMITALL, applied as the first step in the search, takes the initial "bite" of the file. This set is then restricted to documents indexed by the Junior College Clearinghouse. The remaining facets are incorporated into the strategy and the search terminated when the final set is satisfactory to the searcher.

Formulation #1	Comments
LIMITALL/ED	
1 S CH=JC	All ERIC reports indexed by the Junior Colleges Clearinghouse form a facet for JUNIOR COLLEGES.
2 S ENROLLMENT?	
3 C 1 AND 2	At this point 140 citations are retrieved and the searcher continues, adding the third facet to the search.
4 S PROJECTION?	
5 C 3 AND 4	The search is terminated after the third facet is imposed; the final set contains 21 citations.

Formulation #2 employs Search Save #9 to build the JUNIOR COLLEGES facet. As a high recall formulation, it is an example of the building block approach.

Formulation #2	Comments
LIMITALL/EJ	
1 .EXECUTE _____ (Search Save for JUNIOR COLLEGES)	
2 S ENROLLMENT?	Searching is not confined to the descriptor field but encompasses all subject-conveying fields of the ERIC Record.
3 S PROJECT?	
4 S FUTUR?	
5 S ESTIMAT?	
6 S FORECAST?	
7 C 3-6/OR	
8 C 1 AND 2 AND 7	

Step 8: Evaluating Final Results

The answer set to the topic is not available in the ONTAP file; the list of relevant citations follows this discussion. You can figure recall and precision percentages using the formulas summarized below in the evaluation of formulations #1 and #2.

If high recall is a search objective, the second formulation is preferred. Even though all 26 relevant citations are retrieved, the ratio of non-relevant citations in the output to relevant citations is nearly 2 to 1. In order to capture all relevant citations in the ONTAP file, as many search terms as possible

are used to represent the facets JUNIOR COLLEGES and PROJECTIONS. This, in addition to free text searching, may account for the low precision of this formulation.

In contrast, formulation #1, almost a Briefsearch considering its ease and efficiency, achieves high precision and would be favored in cases requiring a high precision, low cost formulation.

Formulation #1

Number of citations found = 21

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{16}{26}$ (62% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{16}{21}$ (76% of retrieved documents are relevant)

Formulation #2

Number of citations found = 68

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{26}{26}$ (100% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{26}{68}$ (38% of retrieved documents are relevant)

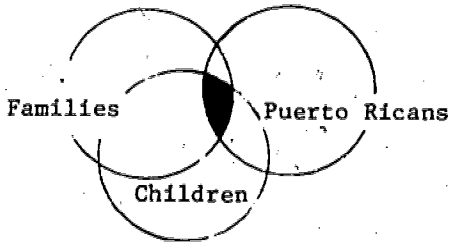
Relevant citations:

ED 110153	ED 110152	ED 110122	ED 110115	ED 110107	ED 109582
ED 105958	ED 105919	ED 105918	ED 105913	ED 104213	ED 103065
ED 101792	ED 101786	ED 100431	ED 100250	ED 099082	ED 099080
ED 099062	098887	ED 098881	ED 097035	ED 096895	ED 095971
ED 095948	ED 095782				

Search topic #14: The Family background of Puerto Rican children

Step 3: Formulating Basic Search Logic -- Planning Search Strategies

Three concepts are present in the search topic: Puerto Ricans, children and families. The diagram below illustrates the search logic:



Step 4: Compiling the Search Terms

Searching vocabulary can be compiled from the ERIC Thesaurus. Below are listed the three facets and terms which are suggested for inclusion in a formulation.

PUERTO RICANS	FAMILIES	CHILDREN
Puerto Rican -s	Family -ies	Child -ren
Puerto Rico	Household -s	Boy -s
	Mother -s	Girl -s
	Father -s	Juvenile -s
	Parent -s	Teenage -r -rs
		Youth -s
		Baby -ies
		Infant
		Adolescent -ce -s
		Toddler -s

Step 5: Ordering Output

The suggested method of approaching this topic is beginning with the most specific facet, PUERTO RICANS, which is a named object and easy to represent,

and adding the other facets to the formulation until the results are satisfactory to the searcher. This mode of search strategy is actually comprised of two approaches, most specific facet first and successive fractions. The building block approach could also be employed as an approach to search strategy; both approaches are illustrated below.

Step 6: Conceptualizing the Search as Input to the Retrieval System

Formulation #1 is a combination of the most specific facet first and successive fractions approaches. The subsearch for PUERTO RICANS is executed first so that the searcher can find out how many items are in the file containing variant forms of Puerto Ricans before formulating another subsearch. Since the set for PUERTO RICANS is considered too large, the searcher continues, evaluating preliminary results after incorporating additional facets to the formulation.

Formulation #1	Comments
1 S PUERTO(W)RICO	
2 S PUERTO(W)RICAN	
3 S PUERTO(W)RICANS	
4 C 1-3/OR	161 citations are retrieved at this point, and the searcher continues rather than terminating the search.
5 .EXECUTE_____(Search Save for for CHILDREN)	
6 C 4 AND 5	The search could be ended here; 54 citations would comprise the output. However, the searcher imposes the third and last facet of the topic, FAMILIES.
7 S FAMILY	
8 S FAMILIES	
9 S HOUSEHOLD?	
10 S PARENT?	
11 S MOTHER?	
12 S FATHER?	
13 C 7-13/OR	
14 C 6 AND 13	The output contains 23 citations.

The Briefsearch contains the terminology used in the search topic: Puerto Ricans, families and children. Since all three facets are employed, it represents the building block approach.

Briefsearch

1 S PUERTO(W)RICANS(C)CHILDREN
2 S FAMIL?
3 C 1 AND 2

Step 8: Evaluating Final Results

Formulation #1 achieves high recall, retrieving all seven relevant items from the ONTAP file. However, since the information seeker must wade through much non-relevant material, the Briefsearch is favored when high precision is a search objective.

Formulation #1

Number of citations found = 23

Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{7}{7}$ (100% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{7}{23}$ (30% of retrieved documents are relevant)

Briefsearch

Number of citations found = 5

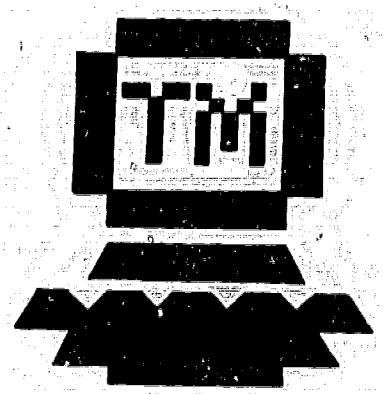
Recall = $\frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}}$ $\frac{3}{7}$ (43% relevant documents retrieved)

Precision = $\frac{\text{Number of relevant citations found}}{\text{Number of citations found}}$ $\frac{3}{5}$ (60% of retrieved documents are relevant)

Relevant citations:

ED 107742 ED 106383 ED 105055 ED 103739 ED 103381 ED 103371
ED 099276

TESTS, MEASUREMENT, AND EVALUATION

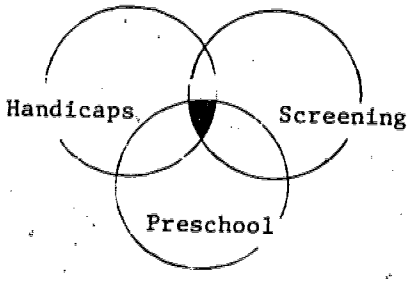


Search topic #15: Screening preschool children for potential learning disabilities or mental handicaps. Items discussing screening for physical handicaps are not wanted.

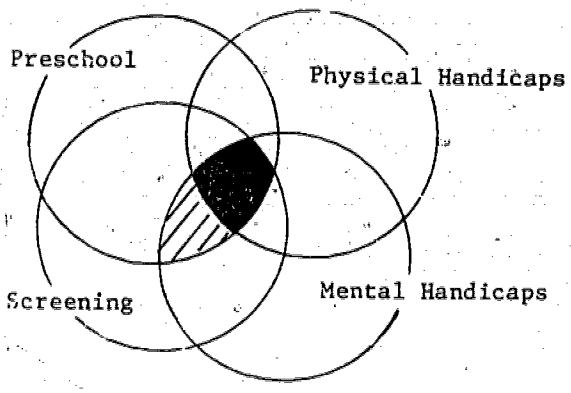
Search Objective: Only journal articles are wanted; also, the information seeker desires high precision and a low cost search.

Step 3: Formulating Basic Logic -- Planning Search Strategies

There are three concepts in the search topic: screening, preschool and handicaps. The portion of the data base intended for retrieval must contain all three facets, so that the three facets are joined by the AND operator in the search logic which is illustrated below.



Since information on physical handicaps is not wanted, the search logic may be represented in another way. In this case, PHYSICAL HANDICAPS forms a fourth facet. The blackened portion in the diagram below contains the facets PRESCHOOL, SCREENING and PHYSICAL HANDICAPS. Since this combination is NOT desired, it is partitioned from the retrieved set using the NOT operator. Rather, the shaded area containing PRESCHOOL, SCREENING and MENTAL HANDICAPS, NOT PHYSICAL HANDICAPS is desired.



Step 4: Compiling the Search Terms

The ERIC Thesaurus is a rich source for search terms representing all three facets. Since two search facets make use of the Search Saves in the appendix, it is evident that this topic is made up of concept groups which occur frequently in ERIC searches. Check the Search Saves for PRESCHOOL and MENTALLY HANDICAPPED as you compile search terms for your own formulation in order to gather additional vocabulary.

Step 5: Ordering Output

Since all three facets are heavily posted, it is expected that all facets will be introduced into the formulation. The building block approach may be a suitable mode for formulating the search, as subsearches for every facet are developed and combined into the formulation. Since low cost is a search objective, the building block approach may not be a wise choice as the combination of all facets may result in no or so few postings that the searcher may have to "backtrack" or reformulate the search; consequently, the searcher may decide to begin the online interaction with a Briefsearch. Employing all facets in the Briefsearch, the searcher may then elect, depending upon the results to follow up with a building block approach. The successive fractions approach is an alternative if the Briefsearch results in few postings.

Another means of reducing online connect time is to apply the LIMITALL command as the first statement in the formulation; in this way, one of the output specifications, i.e., restricting the output to journal articles, is satisfied and the remainder of the search formulation is performed in the appropriate subfile, reducing processing time.

High precision is also a search objective. Adding the fourth facet, PHYSICAL HANDICAPS, to the formulation as a concept not desired may reduce the volume of non-relevant items in the output. Favoring controlled term and title searching instead of free text may also result in fewer non-relevant items in the retrieved set.

It is important that you, the searcher, keep in mind the search objectives stated by the information seeker when selecting an approach to search strategy. In Step 6, four formulations are covered; when you have reviewed each one, select the one you feel satisfies the requestor's needs or develop your own formulation to achieve the search objectives.

Step 6: Conceptualizing the Search in the Language of the Retrieval System

The Briefsearch is used initially in order to determine whether to develop more complex formulations in the building block or successive fractions approach. All three facets are present in the following formulation:

Briefsearch

```
  LIMITALL/EJ
1 S EARLY(W)CHILDHOOD(C)TESTS
2 S HANDICAP?
3 C 1 AND 2
```

Eighteen citations are retrieved by the Briefsearch. The searcher engages the building block approach in the next example. High recall is a primary goal in this case, as two Search Saves are employed to represent the HANDICAPS and

and PRESCHOOL facets and free text searching is used throughout the formulation.

Formulation #2	Comments
LIMITALL/EJ	
1 .EXECUTE _____	(Search Save for PRESCHOOL)
2 .EXECUTE _____	(Search Save for MENTALLY HANDICAPPED)
3 S TEST? ?	Statements 3-8 form the facet for SCREENING
4 S TESTED	
5 S TESTING	
6 S IDENTIF?	
7 S SCREEN?	
8 S DIAGNOS?	
9 C 3-8/OR	
10 C 1 AND 2 AND 9	At this point, 62 citations are retrieved; it is decided to incorporate the fourth facet, PHYSICAL HANDICAPS, to the formulation, in order to exclude items in the output which treat physical handicaps, a topic not wanted.
11 S PHYSICAL?	
12 C 2 NOT 11	In statement 12, items containing the root Physical- are singled out and discarded from the set on MENTAL HANDICAPS.
13 C 12 AND 2 AND 9	Fifty-nine citations are retrieved and the searcher has decided to end the formulation at this point.

Instead of using the comprehensive Search Saves, both formulations #3 and #4 incorporate the terminology used by the information seeker to express the search topic. Only the search terms, Test, its variant forms, and Early Childhood are added to the searching vocabulary in order to increase recall. In formulation #3, free text searching is used.

Formulation #3

LIMITALL/EJ
1 S EARLY(W)CHILDHOOD
2 S PRESCHOOL?
3 S LEARNING(W)DISABILITIES

Formulation #3 (con't.)	Comments
4 S HANDICAP?	
5 S TEST?	Since truncation is used freely with such "common" roots (as opposed to unique) as Test- and Screen-, it can be expected that false drops will occur and affect precision adversely.
6 S SCREEN?	
7 C (1 OR 2) AND (3 OR 4) AND (5 OR 6)	

Thirty citations are retrieved in formulation #3. Using the same search terms, formulation #4 is limited to searching the controlled term fields and the title. It is anticipated that few non-relevant items will be retrieved by this formulation.

Formulation #4	Comments
LIMITALL/EJ	
1 S EARLY(W)CHILDHOOD/TI,ID,DE	
2 S PRESCHOOL/TI,ID,DE	
3 S LEARNING(W)DISABILITIES/TI,ID,DE	
4 S HANDICAPPED/TI,ID,DE	
5 S TESTS/TI,ID,DE	
6 S SCREENING/TI,ID,DE	
7 C (1 OR 2) AND (3 OR 4) AND (5 OR 6)	Twenty-two citations are retrieved at this point.

(Reader: Remembering the search objectives, high precision and low cost, which formulation of those given above would you choose as satisfying these objectives? Have you developed a better strategy to meet the objectives?_____)

Step 8: Evaluating Final Results

Twenty-eight ERIC documents are judged as being relevant to the search topic; the list of these 28 citations are not available in the ONTAP file but are cited by ERIC accession number at the conclusion of this discussion.

Formulation #2 in which high recall is the goal retrieves all the relevant items in the ONTAP file; however, nearly one out of every two citations is

non-relevant. Moreover, the formulation is costly as two lengthy Search Saves must be processed. The results of the search are summarized below:

Formulation #2

Number of citations found = 59

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{28}{28} \quad \text{(100\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{28}{59} \quad \text{(44\% of retrieved documents are relevant)}$$

The purpose of the Briefsearch, in this case, is to perform an efficient and rapid survey of the file in order to determine the appropriate approach to search strategy. It functions adequately, in this role; it could also be employed as the first portion of a citation pearl growing approach in which a few relevant citations are needed in order to gather additional search terms. The formulation itself, as three search statements, recalls nearly half the relevant items in the file. Its recall and precision ratios are given below:

Briefsearch

Number of citations found = 18

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{12}{28} \quad \text{(43\% relevant documents retrieved)}$$

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{12}{18} \quad \text{(66\% of retrieved documents are relevant)}$$

Both formulations #3 and #4 require a minimum of online connect time. They are comprised of heavily posted search terms/descriptors. Precision varies little between the searches but nearly 20% more relevant items are retrieved by formulation #3 which employs free text searching.

Formulation #3. (Free text searching)

Number of citations found = 30

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{21}{28}$$ (75% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{21}{30}$$ (70% of retrieved documents are relevant)

Formulation #4. (Controlled vocabulary and title searching)

Number of citations found = 22

$$\text{Recall} = \frac{\text{Number of relevant citations found}}{\text{Number of citations in answer set}} \quad \frac{16}{28}$$ (57% relevant documents retrieved)

$$\text{Precision} = \frac{\text{Number of relevant citations found}}{\text{Number of citations found}} \quad \frac{16}{22}$$ (73% of retrieved documents are relevant)

Formulation #4 achieves the highest precision; however, formulation #3 results in nearly the same precision with an improvement in recall. In the Briefsearch, two out of three retrieved items are relevant, and it is a low cost formulation as little online connect time is necessary.

Keeping in mind the search objectives, have you developed a formulation which satisfies these objectives? Three of the four formulations given above may satisfy the requestor's needs; which formulation including your own do you prefer? Determining the best formulation is left to you, the reader and searcher.

Relevant citations:

EJ 120892	EJ 119492	EJ 119474	EJ 119468	EJ 119344	EJ 119342
EJ 119253	EJ 115941	EJ 112580	EJ 112578	EJ 112577	EJ 112576
EJ 111026	EJ 109399	EJ 109374	EJ 109371	EJ 109369	EJ 109368
EJ 109367	EJ 107426	EJ 107387	EJ 105970	EJ 104675	EJ 104296
EJ 103429	EJ 102662	EJ 102642	EJ 102322		

APPENDIX

COMMON SEARCH FACETS/SEARCH SAVE FORMULATIONS

When searching the ERIC file, a number of concepts occur frequently as facets of search topics such as

1. Grade school level
2. Population groups
3. Types of libraries
4. Geographic areas
5. Miscellaneous topics

This appendix contains 13 search formulations covering three of the five general areas listed above.* The formulations have been constructed according to the Building Block approach to search strategy; in almost every topic, a single concept is presented. Search logic is fairly simple; the concept, its variant forms and synonymous terms are individually SELECTed and all results COMBINEd by the OR operator. High recall is the primary objective.

Since search terms have been limited to those having postings, it is possible that some terms have been overlooked which may have postings in the future; the formulations can be modified to meet specific needs or used as given. It is suggested that searchers store any number of these searches by using the DIALOG Search Save feature which is illustrated below in a search for Junior High School (grades 7-9).

SAMPLE SEARCH FACET: Junior High School	COMMENTS AND DESCRIPTORS/Identifiers
1 S GRADE 7	GRADE 7
2 S GRADE 8	GRADE 8
3 S GRADE 9	GRADE 9

*All 13 Search Saves in this section and the Search Save for WOMEN (Section II, Search topic #1) have been selected with permission from Selected Search Subroutines for Searching the ERIC Data Base with the Lockheed Information Retrieval Service, by Charles P. Bourne, Barbara Anderson and Jo Robinson. Berkeley, CA: Institute of Library Research, University of California, 1977. (ED 140870)

SAMPLE SEARCH FACET (cont.)	COMMENTS AND DECRIPTORS/Identifiers
4 S JUNIOR(W)HIGH	JUNIOR HIGH SCHOOL ROLE, JUNIOR HIGH SCHOOL STUDENTS, JUNIOR HIGH SCHOOLS/3 multi-term Junior High Identifiers
5 S MIDDLE(W)SCHOOL	
6 S MIDDLE(W)SCHOOLS	MIDDLE SCHOOLS
7 C 1-6/OR	
END/SAVE SERIAL#1JHU 23MAY78 8:52:56 USER 4111 $0.00 0.046 HRS FILE201 13 DESCRIPTORS $0.37 TYMNET $0.37 ESTIMATED TOTAL COST	After typing the search statements and COMBINing them in an OR relationship, the searcher types END/SAVE and concludes the Search Save. The routine is locked into the system when the system assigns a number to the strategy and prints it as part of its response to the END/SAVE command. In this case, serial number 1JHU is assigned. To recall and process the strategy, type .EXECUTE 1JHU

Comments concerning the search formulations are given alongside the strategy as demonstrated above; also mentioned are descriptors (in CAPITAL letters) and identifiers included in the formulation according to the Thesaurus of ERIC Descriptors, 7th edition, 1977, and the ERIC Identifier Usage Report, 1974. The approximate processing time and cost for EXECUTing a Search Save is given at the end of every formulation.

Portions of Search Saves can be executed using the EXECUTE STEPS command. In this way, separate sets for every search statement are created rather than a single set for the final statement of the Search Save. Using the above example, the following example demonstrates how the EXECUTE STEPS function is employed to perform statements 1 to 3 of the Search Save.

SAMPLE SEARCH FACET (cont.)	COMMENTS
.EXECUTE STEPS 1JHU/3	In this example, only statements 1 to 3 are performed. The operation cannot begin after the first statement of the Search Save. If the command does not contain the last step number, e.g. .EXECUTE STEPS 1JHU, the entire Search Save is performed and intermediary results obtained.
1 91 S GRADE 1	
2 103 S GRADE 2	
3 72 S GRADE 3	

At the conclusion of the EXECUTE STEPS command, date and cost information are not given. The system's response looks the same as if the searcher had SELECTed the search phrases and terms manually and not employed the EXECUTE STEPS function.

Search Saves can be exchanged across DIALOG passwords.* A second user can execute the full Search Save on Junior High Schools in the following way:

.EXECUTE 1JHU/USER 4111

If the second user wishes that intermediary sets be created from the exchanged Search Save, the Save is performed in the following way:

.EXECUTE STEPS 1JHU/USER 4111

If any readers develop such Search Saves, we at ERIC/IR would be glad to serve as a clearinghouse for them. Just send us a copy and we will put them from time to time in the ERIC/IR newsletter.

*For a more detailed description of exchanging Search Saves, see "What's New on the System?" Chronolog 6, 2 (Feb. 1978): 4.

PRIMARY EDUCATION

Search Facet #1: Primary education (grades K-3)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S EARLY(W)CHILDHOOD	EARLY CHILDHOOD, EARLY CHILDHOOD EDUCATION/ 13 multi-term Early Childhood Identifiers
2 S GRADE 1	Entered as assigned index terms in order to lessen possibility of false drops; included descriptor, GRADE 1
3 S GRADE 2	GRADE 2
4 S GRADE 3	GRADE 3
5 S KINDERGART?	KINDERGARTEN, KINDERGARTEN CHILDREN/ 7 multi-term Kindergarten Identifiers
6 S PRIMARY(W)EDUCATION	PRIMARY EDUCATION/Primary Education Project
7 S PRIMARY(W)GRADE	Since the truncation feature is not currently operational in statements using word proximity, PRIMARY(W)GRADE and its plural PRIMARY(W)GRADES must be entered in two steps
8 S PRIMARY(W)GRADES	PRIMARY GRADES
9 S PRIMARY(W)PROGRAM	
10 S PRIMARY(W)PROGRAMS	UNGRADED PRIMARY PROGRAMS
11 S PRIMARY(W)SCHOOL	
12 S PRIMARY(W)SCHOOLS	
13 S PRIMARY(W)SYSTEM	NONGRADED PRIMARY SYSTEM
14 S PRIMARY(W)SYSTEMS	
15 C 1-14/OR	

Approximate Processing Time: .075 hrs. ($1.88) in file 1.

ELEMENTARY EDUCATION

Search Facet #2 : Elementary Education (grades K-8)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S EARLY(W)CHILDHOOD	EARLY CHILDHOOD, EARLY CHILDHOOD EDUCATION, 14 multi-term Early Childhood identifiers
2 S ELEMENTARY(1W)EDUCATION	ELEMENTARY EDUCATION, ELEMENTARY SECONDARY EDUCATION/Elementary Education Voucher Demonstration, 16 multi-term variations of Elementary Secondary Education Act
3 S ELEMENTARY(W)GRADE	
4 S ELEMENTARY(W)GRADES	ELEMENTARY GRADES
5 S ELEMENTARY(W)PROGRAM	
6 S ELEMENTARY(W)PROGRAMS	UNGRADED ELEMENTARY PROGRAMS/Elementary Programs
7 S ELEMENTARY(W)SCHOOL	ELEMENTARY SCHOOL COUNSELING, E.S. COUNSELORS, E.S. CURRICULUM, E.S. GUIDANCE, E.S. LIBRARIES, E.S. MATHEMATICS, E.S. ROLE, E.S. SCIENCE, E.S. STUDENTS, E.S. SUPERVISORS, E.S. TEACHERS/ Elementary School, 11 multi-term Elementary School identifiers
8 S ELEMENTARY(W)SCHOOLS	ELEMENTARY SCHOOLS, BIRACIAL ELEMENTARY SCHOOLS, CATHOLIC ELEMENTARY SCHOOLS
9 S GRADE 1	Word proximity, e.g. GRADE(W)1, is not used when searching for grades because the possibility exists for retrieving a substantial number of false drops; GRADE 1
10 S GRADE 2	GRADE 2
11 S GRADE 3	GRADE 3
12 S GRADE 4	GRADE 4

Search Facet #2 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
13 S GRADE 5	GRADE 5
14 S GRADE 6	GRADE 6
15 S GRADE 7	GRADE 7
16 S GRADE 8	GRADE 8
17 S INTERMEDIATE(W)GRADE	
18 S INTERMEDIATE(W)GRADES	INTERMEDIATE GRADES
19 S INTERMEDIATE(W)SCHOOL	Intermediate School 201
20 S INTERMEDIATE(W)SCHOOLS	Intermediate Schools
21 S KINDERGART?	KINDERGARTEN, KINDERGARTEN CHILDREN/ Kindergarten, Kindergartner, Kindergartners, 8 multi-term Kindergarter Identifiers
22 S PRIMARY(W)EDUCATION	PRIMARY EDUCATION/Primary Education Project
23 S PRIMARY(W)GRADE	
24 S PRIMARY(W)GRADES	PRIMARY GRADES
25 S PRIMARY(W)PROGRAM	
26 S PRIMARY(W)PROGRAMS	UNGRADED PRIMARY PROGRAMS
27 S PRIMARY(W)SCHOOL	
28 S PRIMARY(W)SCHOOLS	
29 S PRIMARY(W)SYSTEM	NONGRADED PRIMARY SYSTEM
30 C 1-15/OR	COMBINing search statements in ranges is limited so the COMBINE operation is repeated three times in this formulation.
31 C 16-29/OR	
32 C 30-31/OR	

Approximate Processing Time: .182 hrs. ($4.55) in file 1.

SECONDARY EDUCATION

Search Facet #3: Secondary Education (grades 9-12)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S COLLEGE(W)BOUND	COLLEGE BOUND STUDENTS/College Bound Program, College Bound Programs
2 S COLLEGE(W)PREPARATION	COLLEGE PREPARATION
3 S GRADE 9	GRADE 9
4 S GRADE 10	GRADE 10
5 S GRADE 11	GRADE 11
6 S GRADE 12	GRADE 12
7 S HIGH(W)SCHOOL	12 multi-term HIGH SCHOOL Descriptors, 2 multi-term JUNIOR HIGH SCHOOL Descriptors/ Junior High School, 2 multi-term Junior High School Identifiers
8 S HIGH(W)SCHOOLS	HIGH SCHOOLS, 7 multi-term HIGH SCHOOLS Descriptors, JUNIOR HIGH SCHOOLS
9 S NONCOLLEGE(W)PREPARATORY	NONCOLLEGE PREPARATORY STUDENTS
10 S PRECOLLEGE	Precollege Centers
11 S PRE(W)COLLEGE	Pre College Centers
12 S PREP(W)SCHOOL	
13 S PREPARATORY(W)SCHOOL	
14 S PREPARATORY(W)SCHOOLS	
15 S SECONDARY(W)EDUCATION	SECONDARY EDUCATION, 2 multi-term SECONDARY EDUCATION Descriptors/Secondary Education, Secondary Education Act, Elementary Secondary Education
16 S SECONDARY(W)GRADE	
17 S SECONDARY(W)GRADES	SECONDARY GRADES

Search Facet #3 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
18 S SECONDARY(W)SCHOOL	5 multi-term SECONDARY SCHOOL Descriptors/ Secondary School, 6 multi-term Secondary School Identifiers
19 S SECONDARY(W)SCHOOLS	SECONDARY SCHOOLS, BIRACIAL SECONDARY SCHOOLS
20 C 1-19/OR	

Approximate Processing Time: .170 hrs. ($4.25) in file 1.

CHILDREN

Search Facet #4: Children

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S ADOLESCEN?	ADOLESCENCE, ADOLESCENT LITERATURE, ADOLESCENTS/5 multi-term Adolescent Identifiers
2 S BOY	Boy Girl Identity Task, Boy Scouts, 3 multi-term Boy Identifiers
3 S BOYS	Entered without truncation because of possibility of returned false drops
4 S BOY'S	
5 S CHILD?	CHILD ABUSE, 22 other multi-term CHILD Descriptors, CHILDHOOD ATTITUDES, 5 other multi-term CHILDHOOD Descriptors CHILDREN, 19 other multi-term CHILDREN Descriptors, CHILDRENS ART, CHILDRENS BOOKS, CHILDRENS GAMES, CHILDRENS LITERATURE/41 multi-term Child Identifiers, 2 multi-term Children Identifiers, 13 multi-term Children's Identifiers
6 S GIRL?	GIRLS CLUBS/ 3 multi-term Girl Identifiers
7 S JUVENILE?	JUVENILE COURTS, JUVENILE GANGS, 10 multi-term Juvenile Identifiers
8 S TEENAGE?	TEENAGERS/3 multi-term Tennage Identifiers
9 S YOUTH?	YOUTH, 16 multi-term YOUTH Descriptors/ 29 multi-term Youth Identifiers
10 C 1-9/OR	

Approximate Processing Time: .030 hrs. ($0.75) in file 1.

Search Facet #4: (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
If the search were to cover very young children, then you might want to add the following terms after the search save is EXECUTEd.	
2 S BABY	Baby Talk
3 S BABIES	
4 S INFAN?	INFANCY, INFANT BEHAVIOR, INFANT MORTALITY INFANTS, PREMATURE INFANTS/Infant Care, 19 other multi-term Infant-Identifiers, Infantilization, Infantile Amnesia
5 S TODDLER?	
6 C 1-5/OR	

Approximate Processing Time for Search Save Plus Additional Terms: .074 hrs. ($1.85) in file 1.

AMERICAN INDIANS

Search Facet #5: American Indians (general search)

If the objective is a high recall search, then consideration should be given to formulating a Search Save which includes many specific tribal names. This routine is comprised of general terms applied to the large population group of American Indians and does not cover individual tribal names.

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S ALASKA(W)NATIVE	Alaska Native Needs Assessment
2 S ALASKA(W)NATIVES	ALASKA NATIVES
3 S ALASKAN(W)NATIVE	Alaskan Native Needs Assessment
4 S ALASKAN(W)NATIVES	Alaskan Natives
5 S AMERINDIAN?	Amerindians
6 S AMERICAN(W)INDIAN	AMERICAN INDIAN CULTURE, AMERICAN INDIAN LANGUAGES/American Indian, 8 multi-term American Indian Identifiers
7 S AMERICAN(W)INDIANS	AMERICAN INDIANS, NONRESERVATION AMERICAN INDIANS
8 S NATIVE(W)AMERICAN	Native American Administrator Program, Native American Studies Programs
9 S NATIVE(W)AMERICANS	Native Americans
10 S RESERVATIONS (INDIAN)	RESERVATIONS (INDIAN); The descriptor for RESERVATIONS is used here rather than the free text term since RESERVATIONS has a number of meanings, all differing from one another.
11 S TRIBAL	4 multi-term Tribal Identifiers
12 S TRIBE?	TRIBES
13 C 1-12/OR	

Approximate Processing Time: .030 hours ($0.75) in file 1.

AUDIOVISUAL AIDS

Search Facet #6: Audiovisual Aids

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S AUDIO(W)VIDEO	AUDIO VIDEO LABORATORIES/3 multi-term Audio Video Identifiers
2 S AUDIOVIDEO?	
3 S AUDIO(W)VISUAL	9 multi-term Audio Visual Identifiers
4 S AUDIO(W)VISUALS	
5 S AUDIOVISUAL?	7 multi-term AUDIOVISUAL Descriptors/10 multi-term Audiovisual Identifiers
6 S AV	The truncation capability is not used in this search statement because terms such as avenue, aviation, average, or avocation, would be retrieved; included identifier: AV Communication Review
7 S CARTOON?	CARTOONS
8 S CARTRIDG?	
9 S CASSETTE?	3 multi-term CASSETTE or CASSETTES Descriptors/Cassettes, 5 multi-term Cassette Identifiers
10 S CENTRAL(W)SOUND	CENTRAL SOUND SYSTEMS
11 S COLOR(W)PRESENTATION	COLOR PRESENTATION
12 S COLOR(W)PRESENTATIONS	
13 S DIAL(W)ACCESS	DIAL ACCESS INFORMATION SYSTEMS/Dial Access Information Retrieval System, Dial Access Laboratories
14 S DISPLAY(W)PANEL	
15 S DISPLAY(W)PANELS	DISPLAY PANELS
16 S DOCUMENTAR?	DOCUMENTARIES/Documentary Drama, Documentary History, Documentaries

Search Facet #6 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
17 S EXHIBIT?	EXHIBITS
18 S EXPOSITION?	EXPOSITIONS
19 S FELT(W)BOARD	
20 S FELT(W)BOARDS	
21 S FILM?	FILMS, 7 multi-term FILM Descriptors, 4 multi-term FILMS Descriptors, FILMSTRIP PROJECTORS, FILMSTRIPS/31 multi-term Film Identifiers, Filmic Styles, Filmlists, Filmographies, Filmography, Filmography of Films about Movies, Films Incorporated, Filmstrip Viewers
22 S GRAPHIC(W)AID	
23 S GRAPHIC(W)AIDS	
24 S GRAPHIC(W)ART	
25 S GRAPHIC(W)ARTS	GRAPHIC ARTS/Graphic Arts Industries
26 S GRAPHICS	COMPUTER GRAPHICS, ENGINEERING GRAPHICS/ Graphics Expression Reading Improvement System, Graphics Expression System
27 S INSTRUCTIONAL(W)AID	
28 S INSTRUCTIONAL(W)AIDS	INSTRUCTIONAL AIDS/Instructional Aids
29 S INSTRUCTIONAL(W)MEDIA	INSTRUCTIONAL MEDIA/Instructional Media Centers
30 S LANGUAGE(W)AID	
31 S LANGUAGE(W)AIDS	LANGUAGE AIDS
32 S LANGUAGE(W)LAB	
33 S LANGUAGE(W)LABS	
34 S LANGUAGE(W)LABORATORY	LANGUAGE LABORATORY EQUIPMENT, LANGUAGE LABORATORY USE/Language Laboratory, Language Laboratory Directors, Language Laboratory Equipment

Search Facet #6 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
35 S LANGUAGE(W)LABORATORIES	LANGUAGE LABORATORIES
36 S LEARNING(W)LABORATORY	Learning Laboratory
37 S LEARNING(W)LABORATORIES	LEARNING LABORATORIES
38 S MICROPHONE?	MICROPHONES
39 S MULTI(W)MEDIA	Multi Media Techniques, Multi-Media Treatise...
40 S MULTIMEDIA	MULTIMEDIA INSTRUCTION/5 multi-term Multimedia Identifiers
41 S OVERHEAD(W)PROJECTOR	Overhead Projector Animation System
42 S OVERHEAD(W)PROJECTORS	OVERHEAD PROJECTORS
43 S PHONODISC?	
44 S PHONOGRAPH?	PHONOGRAPH RECORDS, LANGUAGE RECORDS (PHONOGRAPH)/Phonographic Representation, Phonographs
45 S PHONOTAPE	PHONOTAPE RECORDINGS/Phonotape Records
46 S PHOTOGRAPH?	PHOTOGRAPHIC EQUIPMENT, PHOTOGRAPHS, PHOTOGRAPHY/2 multi-term Photograph Identifiers, 3 multi-term Photographic Identifiers
47 S POSTER?	
48 S PROJECTION(W)EQUIPMENT	PROJECTION EQUIPMENT
49 S RADIO	RADIO, EDUCATIONAL RADIO, RADIO TECHNOLOGY/ 17 multi-term Radio Identifiers
50 S RADIOPHONIC	Radiophonic Teaching
51 S RADIOS	
52 S REALIA?	REALIA
53 S SCREENS (DISPLAYS)	SCREENS (DISPLAYS)
54 S SLIDE(W)PROJECTOR	
55 S SLIDE(W)PROJECTORS	

Search Facet #6 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
56 S SLIDE(W)TAPE	Slide Tape Sequences
57 S SLIDE(W)TAPES	
58 S SLIDES	SLIDES
59 S TAPE(W)RECORDER	
60 S TAPE(W)RECORDERS	TAPE RECORDERS
61 S TAPE(W)RECORDING	
62 S TAPE(W)RECORDINGS	TAPE RECORDINGS, VIDEO TAPE RECORDINGS
63 S TELEVISION?	TELEVISION, 21 multi-term TELEVISION Descriptors/16 multi-term Television Identifiers
64 S TRANSPARENC?	TRANSPARENCIES
65 S TV	Identifiers: TV, TV College, TV College (Chicago)
66 S TV'S	
67 S VIDEO?	4 multi-term VIDEO Descriptors/11 multi-term Video Identifiers, Videographics Systems, Videopublishing, Videorecords, Videosonic Opticon Application, Videosonic Teaching Machine, Videotape Participation System, Videotape Program Service
68 S VIEWER?	
69 S VIEWING(W)CENTER	
70 S VIEWING(W)CENTERS	
71 S VISUAL(W)AID	
72 S VISUAL(W)AIDS	VISUAL AIDS/Visual Aids

Search Facet #6 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
73 S VISUALS	
74 C 1-25/OR	
75 C 26-50/OR	
76 C 51-73/OR	
77 C 74-76/OR	

Approximate Processing Time: .261 hours. ($6.53) in file 1.

EVALUATION (HIGH RECALL)

Search Facet #7: EVALUATION

This is a high recall search for the concept of Evaluation. Full use of free text searching and truncation is made; over 150,000 postings are retrieved when processing this search in DIALOG's full ERIC. Search Save #8, on the other hand, is a high precision formulation for Evaluation.

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S ANALY?	30 Multi-term ANALYSIS Descriptors, ANALYSTS ANALYTIC GEOMETRY, ANALYTICAL CRITICISM/ 5 multi-term Analysis Identifiers 8 multi-term Analytical Identifiers, Analyticity, Analytical Methods, Analytic Philosophy, Analytic Rotational Methods
2 S APPRAIS?	PROPERTY APPRAISAL/Appraisal of Personal Growth
3 S ASSESS?	ASSESSED VALUATION, EDUCATIONAL ASSESSMENT, INFORMAL ASSESSMENT, NEEDS ASSESSMENT, PERSONALITY ASSESSMENT/3 multi-term Assessing Identifiers, Assessment, 7 multi-term Assessment Identifiers
4 S COMPAR?	COMPARATIVE ANALYSIS, COMPARATIVE EDUCATION COMPARATIVE STATISTICS, COMPARATIVE TESTING, AUDIO ACTIVE COMPARE LABORATORIES, EDUCATIONAL STATUS COMPARISON/Comparability, Comparable Cost Information, 23 multi-term Comparative Identifiers, Comparing Reading Approaches
5 S COST	COST EFFECTIVENESS, COST INDEXES/Cost Accounting, 9 multi-term Cost Identifiers
6 S COSTS	COSTS, 9 multi-term COSTS Descriptors/ Costs of Schools Training and Education

Search Facet #7 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
7 S CRITER?	CRITERIA, 6 multi-term CRITERIA Descriptors, CRITERION REFERENCED TESTS/Criteria of Success In English Test, 5 multi-term Criterion Identifiers
8 S EFFECT?	PRIMACY EFFECT, EFFECTIVE TEACHING, 4 multi-term EFFECTIVENESS Descriptors, 4 multi-term EFFECTS Descriptors/Effectance Motivation, Effectively Influencing Political Decisions, Effectiveness, 5 multi-term Effective Identifiers, Effects of Daylight Savings Time.
9 S EVALUAT?	EVALUATION, ADMINISTRATOR EVALUATION, 26 other multi-term EVALUATION Descriptors, EVALUATIVE THINKING/Evaluation, 10 multi-term Evaluation Identifiers, 7 multi-term Evaluative Identifiers, Evaluators
10 S FEASIB?	FEASIBILITY STUDIES
11 S IMPACT?	IMPACT, 7 multi-term Impact Identifiers
12 S MEASUR?	MEASUREMENT, 6 multi-term MEASUREMENT Descriptors, VISUAL MEASURES/6 multi-term Measurement Identifiers, Measure of Sampling Adequacy, 4 multi-term Measures Identifiers, Measuring Instruments
13 S PERFORMANCE?	PERFORMANCE, 11 multi-term PERFORMANCE Descriptors/19 multi-term Performance Identifiers

Search Facet #7 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
14 S RELIAB?	RELIABILITY, TEST RELIABILITY/Reliability Analysis Center
15 S SUCCESS	SUCCESS FACTORS
16 S SUCCESSES	
17 S SUCCESSFUL	
18 S VALID?	VALIDITY, PREDICTIVE VALIDITY, TEST VALIDITY/Valid Confidence, Validity Research, Validity Study Service
19 C 1-8/OR	

Approximate Processing Time: .068 hrs. ($1.70) in file 1.

EVALUATION (HIGH PRECISION)

Search Facet #8. EVALUATION

This is a high precision search formulation for the concept of Evaluation. It is restricted to searching only the identifier, title, and descriptor fields of the ERIC record. Truncation is not used at all in this formulation in order to eliminate the possibility of retrieving false drops. Because searching is restricted to three of the 7 possible subject-conveying fields in the ERIC record, the result of Search Save #8 is only half the size of the final set in #7.

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S ANALYSIS/DE,ID,TI	30 multi-term ANALYSIS Descriptors/5 multi-term Analysis Identifiers
2 S ANALYSES/DE,ID,TI	
3 S ANALYTIC/DE,ID,TI	ANALYTIC GEOMETRY
4 S APPRAISAL/DE,ID,TI	PROPERTY APPRAISAL/Appraisal of Personal Growth in Teamwork
5 S APPRAISING/DE,ID,TI	
6 S ASSESSING/DE,ID,TI	3 multi-term Assessing Identifiers
7 S ASSESSMENT/DE,ID,TI	EDUCATIONAL ASSESSMENT, INFORMAL ASSESSMENT, NEEDS ASSESSMENT, PERSONALITY ASSESSMENT/ Assessment, 7 multi-term Assessment Identifiers
8 S COMPARATIVE/DE,ID,TI	4 multi-term COMPARATIVE Descriptors/23 multi-term Comparative Identifiers
9 S COMPARING/DE,ID,TI	2 multi-term Comparing Identifiers
10 S COMPARISON/DE,ID,TI	
11 S COSTS/DE,ID,TI	COSTS, 9 multi-term COSTS Descriptors/ Costs of Schools Training and Education

Search Facet #8 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
12 S CRITERIA/DE,ID,TI	CRITERIA, 6 multi-term CRITERIA Descriptors/ Criteria of Success in English Test
13 S CRITERION/DE,ID,TI	CRITERION REFERENCED TESTS/5 multi-term Criterion Identifiers
14 S EFFECTIVENESS/DE,ID,TI	4 multi-term EFFECTIVENESS Descriptors
15 S EVALUATION/DE,ID,TI	EVALUATION, 26 multi-term EVALUATION Descriptors/Evaluation, 10 multi-term Evaluation Identifiers
16 S EVALUATIVE/DE,ID,TI	EVALUATIVE THINKING/8 multi-term Evaluative Identifiers
17 S FEASIBILITY/DE,ID,TI	FEASIBILITY STUDIES
18 S IMPACT/DE,ID,TI	
19 S MEASUREMENT/DE,ID,TI	MEASUREMENT, 6 multi-term MEASUREMENT Descriptors/6 multi-term Identifiers
20 S PERFORMANCE/DE,ID,TI	PERFORMANCE, 11 multi-term PERFORMANCE Descriptors/19 multi-term Performance Identifiers
21 S RELIABILITY/DE,ID,TI	RELIABILITY, TEST RELIABILITY/Reliability Analysis Center
22 S SUCCESSFUL/DE,ID,TI	
23 S VALID/DE,ID,TI	Valid Confidence
24 S VALIDITY/DE,ID,TI	VALIDITY, PREDICTIVE VALIDITY, TEST VALIDITY/Validity Research, Validity Study Service
25 C 1-24/OR	

Approximate Processing Time: .031 hrs. ($.78) in file 1.

JUNIOR COLLEGE

Search Facet #9: Junior College

The terms JUNIOR COLLEGE and COMMUNITY COLLEGE are searchable in all subject-conveying fields except corporate source, sponsoring agency and descriptive note in order to exclude items sponsored and/or authored by a specific educational institution.

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S ASSOCIATE(W)DEGREE	
2 S ASSOCIATE(W)DEGREES	ASSOCIATE DEGREES
3 S COMMUNITY(W)COLLEGE/DE, ID,TI,AB	8 multi-term Community College Identifiers
4 S COMMUNITY(W)COLLEGES	COMMUNITY COLLEGES
5 S GRADE 13	GRADE 13
6 S GRADE 14	GRADE 14
7 S JUNIOR(W)COLLEGE/DE,ID, TI,AB	JUNIOR COLLEGE LIBRARIES, JUNIOR COLLEGE STUDENTS/6 multi-term Junior College Identifiers
8 S JUNIOR(W)COLLEGES	JUNIOR COLLEGES
9 S POST(W)SECONDARY	POST SECONDARY EDUCATION/Post Secondary Schools
10 S POSTSECONDARY	
11 C 1-10/OR	

Approximate Processing Time: .049 hrs. ($1.23) in file 1.

HIGHER EDUCATION

Search Facet #10: Higher Education

The terms COLLEGE, COLLEGES, HIGHER EDUCATION, GRADUATE SCHOOL, UNIVERSITY, and UNIVERSITIES are searchable in all subject-conveying fields except corporate source, sponsoring agency and descriptive note fields. This reduces the number of false drops which may occur when a particular educational institution is responsible for the document as a corporate author or sponsor.

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S COLLEGE/DE,ID,TI,AB	COLLEGE ADMINISTRATION, 40 other multi-term COLLEGE Descriptors/College, over 60 multi-term College Identifiers
2 S COLLEGES/DE,ID,TI,AB	COLLEGES, 13 multi-term COLLEGES Descriptors
3 S COLLEGIA?	Collegial System, Collegial Teams, four multi-term Collegiate Identifiers
4 S DOCTORAL	DOCTORAL DEGREES, POST DOCTORAL EDUCATION, DOCTORAL PROGRAMS, DOCTORAL THESES/ Doctoral These, Doctoral Thesis
5 S GRADUATE(W)SCHOOL/DE,ID,TI,AB	Graduate School Foreign Language Test, Graduate School of Education
6 S GRADUATE(W)SCHOOLS	Graduate Schools, 2 multi-term Graduate Schools Identifiers
7 S GRADUATE(W)STUDENT 8 S GRADUATE(W)STUDENTS	GRADUATE STUDENTS
9 S GRADUATE(W)STUDY 10 S GRADUATE(W)STUDIES	GRADUATE STUDY
11 S HIGHER(W)EDUCATION/DE,ID,TI,AB	HIGHER EDUCATION/32 multi-term Higher Education Identifiers
12 S POST(W)SECONDARY	POST SECONDARY EDUCATION/Post Secondary Schools
13 S POSTSECONDARY	
14 S UNDERGRADUATE?	UNDERGRADUATE STUDENTS, UNDERGRADUATE STUDY/ 5 multi-term Undergraduate Identifiers

Search Facet #10 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
15 S UNIVERSITY/DE,ID,TI,AB	Three multi-term UNIVERSITY Descriptors/ University, over 100 multi-term Identifiers
16 S UNIVERSITIES/DE,ID,TI,AB	UNIVERSITIES, 3 multi-term UNIVERSITIES Descriptors/3 multi-term Universities Identifiers
17 C 1-16/OR	

Approximate Processing Time: .076 hrs. ($1.90) in file 1.

HANDICAPPED

Search Facet #11: Mentally, neurologically and emotionally handicapped

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S APHASI?	APHASIA/Asphasic
2 S AUTISM	AUTISM
3 S BRAIN(W)DAMAGE	Brain Damage
4 S BRAIN(W)DAMAGED	Brain Damaged
5 S BRAIN(W)DYSFUNCTION	
6 S BRAIN(W)INJURED	MINIMALLY BRAIN INJURED
7 S BRAIN(W)INJURY	
8 S DEVIANT(W)BEHAVIOR	SOCIALLY DEVIANT BEHAVIOR
9 S DYSLEX?	DYSLEXIA
10 S HANDICAP?	HANDICAP DETECTION, HANDICAPPED, 21 multi-term HANDICAPPED DESCRIPTORS, 3 multi-term HANDICAPS DESCRIPTORS
11 S HYPERACTIV?	HYPERACTIVITY
12 S MALADJUSTED(W)CHILD	
13 S MALADJUSTED(W)CHILDREN	
14 S MALADJUSTMENT	MALADJUSTMENT, EMOTIONAL MALADJUSTMENT
15 S MONGOLISM	MONGOLISM
16 S MONGOLOID?	
17 S NEUROS?	NEUROSIS
18 S NEUROTIC?	NEUROTIC CHILDREN/Neuroticism, Neuroticism Scale Questionnaire
19 S PROBLEM(W)CHILD	
20 S PROBLEM(W)CHILDREN	PROBLEM CHILDREN
21 S PSYCHOS?	PSYCHOSIS
22 S PSYCHOTIC?	PSYCHOTIC CHILDREN/Psychotic Impatient Profile
23 S RETARDED	3 multi-term RETARDED Descriptors/ 4 multi-term Retarded Identifiers
24 S SCHIZOPHRENI?	SCHIZOPHRENIA/Schizophrenia
25 C 1-24/OR	
26 S BEHAVIOR?	BEHAVIOR PROBLEMS/Behavior Problem Checklist, Behavioral Problems
27 S EMOTIONAL?	EMOTIONALLY DISTURBED, EMOTIONALLY DISTURBED CHILDREN/Emotionally Disturbed Children
28 S LEARNING	Learning Disability Teaching Consultant, LEARNING DISABILITIES/Learning Disabilities Act of 1969

Search Facet #11 (cont.)

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
29 S MENTAL?	MENTALLY HANDICAPPED, CUSTODIAL MENTALLY HANDICAPPED, EDUCABLE MENTALLY HANDICAPPED, TRAINABLE MENTALLY HANDICAPPED
30 S NEUROLOGICAL?	NEUROLOGICAL DEFECTS/NEUROLOGICALLY HANDICAPPED/Neurologically Handicapped
31 S PERCEPTUAL?	PERCEPTUALLY HANDICAPPED
32 S PERSONALITY	PERSONALITY PROBLEMS
33 C 26-32/OR	
34 S DEFECT?	
35 S DEFICIEN?	
36 S DISABILIT?	
37 S DISORDER?	
38 S DISTURB?	
39 S ILL	
40 S ILLNESS?	
41 S PROBLEM?	
42 C 34-41/OR	
43 C 33 AND 42	
44 C 25 OR 43	

Approximate Processing Time: .139 hrs. ($3.48) in file 1.

PRESCHOOL

Search Facet #12: Preschool

FORMULATION	COMMENTS AND DESCRIPTORS/Identifiers
1 S EARLY(W)CHILDHOOD	EARLY CHILDHOOD, EARLY CHILDHOOD EDUCATION/ 13 multi-term Early Childhood Identifiers
2 S KINDERGART?	KINDERGARTEN, KINDERGARTEN CHILDREN/ 7 multi-term Kindergarten Identifiers
3 S NURSERY(W)SCHOOL	
4 S NURSERY(W)SCHOOLS	Nursery Schools
5 S PREKINDERGART?	
6 S PRE(W)KINDERGARTEN	
7 S PRE(W)PRIMARY	
8 S PREPRIMARY	
9 S PRE(W)SCHOOL	
10 S PRESCHOOL?	PRESCHOOL CHILDREN, 9 multi-term PRESCHOOL DESCRIPTORS/20 multi-term Preschool Identifiers

Approximate Processing Time: .030 hrs. ($0.75) in file 1.

The above formulation is rather specific to the preschool age group. If a high recall search is desired, consideration might be given to adding the following two terms which include the preschool age group, but also include other groups.

11 S CHILD(W)CARE	CHILD CARE, CHILD CARE CENTERS, MIGRANT CHILD CARE CENTERS, CHILD CARE OCCUPATIONS, CHILD CARE WORKERS/Child Care Center
12 S DAY(W)CARE	FAMILY DAY CARE, DAY CARE PROGRAMS, DAY CARE SERVICES/Day Care Centers, 5 multi-term Day Care Identifiers

APPENDIX

GUIDE TO DIALOG'S ERIC DATA BASE

The database description for ERIC prepared by DIALOG for their volume entitled, Guide to DIALOG -- Databases (Palo Alto, CA: Lockheed Information Systems, August 1977) is reproduced here almost in its entirety (only a few pages of Search Examples are omitted).

This description should be an aid to the reader of this manual as it provides the most definitive exposition of the ERIC files (RIE and CIJE) as they appear and are accessed on the DIALOG Information Retrieval Service.

DIALOG* INFORMATION RETRIEVAL SERVICE

ERIC

FILE DESCRIPTION

ERIC is the complete database on educational materials from the Educational Resources Information Center. It consists of two main files: *Resources in Education* (RIE), which is concerned with identifying the most significant and timely education research reports; and *Current Index to Journals in Education* (CIJE), an index of more than 700 publications of interest to every segment of the educational profession. (¶ 1.1)

SUBJECT COVERAGE

The ERIC database includes an enormous variety of educational information in the following general areas: (¶ 1.2)

- Career Education
- Counseling and Personnel Services
- Early Childhood Education
- Educational Management
- Handicapped and Gifted Children
- Higher Education
- Information Resources
- Junior Colleges
- Languages and Linguistics
- Reading and Communication Skills
- Rural Education) and Small Schools
- Science, Mathematics, and Environmental Education
- Social Studies/Social Science Education
- Teacher Education
- Tests, Measurement, and Evaluation
- Urban Education

SOURCES

ERIC indexes a wealth of document types: research reports, evaluation studies, curriculum guides, lesson plans, bibliographies, course descriptions, theses, journal articles, pamphlets and other "fugitive" materials. Most of the items, aside from journal articles, can be purchased from the ERIC Document Reproduction Service (EDRS) in paper copy or microfiche. There are approximately 650 locations throughout the country having collections of the ERIC microfiche, and most are open to the general public. (¶ 1.3)

DIALOG FILE DATA

(¶ 1.4)

Inclusive Dates:	1966 to the present
Update Frequency:	Monthly (approximately 3,000 a month)
File size:	280,000 citations, as of June 1977

ORIGIN

ERIC is produced by: (¶ 1.5)

National Institute of Education
Educational Resources Information Center
Washington, DC 20208

Questions concerning this file should be directed to: (¶ 1.6)

User Services Coordinator
ERIC Processing and Reference Facility
4833 Rugby Avenue, Suite 303
Bethesda, MD 20014

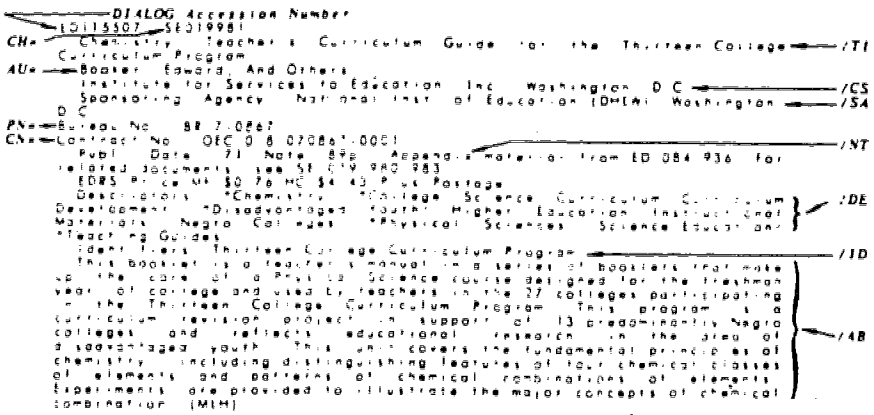
Telephone: 301/656-9723

*Trademark Reg. U.S. Pat. & Trademark Office

FILE 1

ERIC
DIALOG FILE 1

SAMPLE RECORD



RETRIEVAL METHODS

(9.17)

SUBJECT OR TEXT SEARCHING

PAGE	SUFFIX	FIELD NAME	EXAMPLES
1-3	None	Basic Index(includes all seven fields listed below)	E CURRICULUM S CHEMISTRY S SCIENCE(F)EDUCATION
1-3	/AB	Abstract	S NEGRO(W)COLLEGES/AB
1-4	/CS	Corporate Source	S SERVICES(1W)EDUCATION(F)WASHINGTON/CS
1-4	/DE	Descriptor	S PHYSICAL SCIENCES
1-8	/ID	Identifier	S THIRTEEN(W)COLLEGE(W)CURRICULUM/ID
1-10	/NT	Descriptive Note	S APPENDIX/NT
1-10	/SA	Sponsoring Agency	S NATIONAL(W)INST(1W)EDUCATION/SA
1-11	/TI	Title	S CURRICULUM(W)PROGRAM/TI

[1]Also/DE*,/DF,/DF*,/ID*,/IF,/IF*

CODE SEARCHING

PAGE	PREFIX	FIELD NAME	EXAMPLES	
1-11	AC=	Legislative Authority Code	E AC=5b	S AC=95
1-12	AU=	Author	E AU=BOOHER, E	S AU=BOOKER, EDWARD
1-13	CH=	Clearinghouse Code	E CH=VT	S CH=SE
1-15	CN=	Contract/Grant Number	E CN=OEC	S CN=OEC-0-8-07567-0001
1-15	DT=	Document Type	E DT=8	S DT=C
1-16	GC=	Group Code	E GC=110 (CURRICULUM)	(EXPAND only)
1-17	IS=	Issue	E IS=CIJE1977	S IS=RIEAPR76
1-17	JO=	Journal Name	E JO=AAUP	S JO=INTERCOM
1-18	PN=	Bureau/Project Number	E PN=BR	S PN=BR-7-0867
1-18	RN=	Report Number	E RN=B	S RN=8-1897
1-18	SC=	Corporate Source Code or Sponsoring Agency Code	E SC=FGK E SC=BBB	S SC=FGK37812 S SC=BBB06621
1-19	UD=	Update	E UD=7512	S UD=7604
1-19	YR=	Year	E YR=74	S YR=76

LIMITING

PAGE	SUFFIX	FIELD NAME	EXAMPLES
1-19	/ED	Accession Numbers and/or ED Subfile	L1/000788-095253/ED L4/ED
	/EJ	Accession Numbers and/or EJ Subfile	L5/083165-101872/EJ L8/EJ
1-20	/AVAIL	Document Available from EDRS	L6/AVAIL
	/UNAVAIL	Document Not Available from EDRS	L14/UNAVAIL
1-20	/MAJ	Major Descriptor or Identifier	L7/MAJ
	/MIN	Minor Descriptor or Identifier	L9/MIN

DIRECT ACCESS

PAGE	PREFIX	FIELD NAME	EXAMPLES
1-24	None	DIALOG Accession Number	T ED115507/3 D ED115507

FORMATS AVAILABLE

(9.18)

PAGE	NUMBER	RECORD CONTENT
1-22	Format 1	DIALOG Accession Number
	Format 2	Complete Record Except Abstract
	Format 3	Bibliographic Citation
	Format 4	Abstract
	Format 5	Full Record
	Format 6	Title

SEARCH OPTIONS

BASIC INDEX. If a prefix or suffix is not specified in an EXPAND or SELECT command, the command is executed in the basic index. The basic index contains all assigned descriptors (*Thesaurus* terms) and identifiers, plus all meaningful individual words extracted from the Title, Abstract, Descriptor, Identifier, Corporate Source, Descriptive Note, and Sponsoring Agency fields. Thus, all words of subject importance can be searched without regard to their location within a given record. However, the full-text capability can be used to define a search statement further by specifying the desired proximity and order of words, their occurrence in a given field, or by any combination of these as shown in the following sections.

(¶ 1.9)

ABSTRACT (/AB). Abstracts for the report literature (RIE items) are generally short (less than 200-word) summaries of the document contents. Abstracts for the journal articles (CIJE items) are generally brief (less than 50-word) annotations. Annotating is done only to enrich those articles with uninformative, non-expository titles.

(¶ 1.10)

The use of free language terms that might occur in abstracts, titles, or identifiers is particularly useful for search topics in new research areas for which the *Thesaurus* vocabulary may not be adequate.

(¶ 1.11)

Acronyms may be used in abstracts although individual words generally are not abbreviated. The searcher should include the full phrase and its acronym in the strategy.

(¶ 1.12)

```
? SELECT MMPI/AB
          1    130 MMPI/AB
? TYPE 1/4
1
EJ149928
  This study investigated objections to Minnesota Multiphasic
Personality Inventory (MMPI) items by black and white males and
females classified as high or low trust by Rotter's Interpersonal
Trust Scale. Low-trust subjects objected to more items than high-trust
subjects. Race and sex did not influence number of items objected to.
(Author)

? SELECT MINNESOTA(W)MULTIPHASIC(W)PERSONALITY/AB
          2     79 MINNESOTA(W)MULTIPHASIC(W)PERS
? COMBINE 1 OR 2
          3    163  1 OR 2
```

Names of individuals in the text of the abstract may or may not include initials or given name with the surname and may occur in the possessive form. Truncation can be used with the surname to encompass both forms.

(¶ 1.13)

```
? SELECT PIAGET?
          4   1428 PIAGET?
? TYPE 4/4
1
EJ152391
  Describes a study to extend research of Piaget and Jahoda into
children's spatial conceptions. Individual testing and interviewing
```

FILE 1

are discussed. Results indicate differences in rate of spatial stages development of Swiss, Scottish and American children, although Piagets' original theorized stages remain invariate. (Author/AV)

CORPORATE SOURCE (/CS). The Corporate Source field, found only in RIE items, is used to list the organization responsible for the issuance of the document. (It is not the address of the first personal author.) Each corporate source name has been entered in a standardized form and corresponds to a specific eight-character alphanumeric institution code searchable in the SC= field.

(¶ 1.14)

Searching in the /CS field will generally yield broader retrieval than the SC= field because unique codes are assigned to each division of each institution.

(¶ 1.15)

Full-text methods should be used to search corporate source names. It is generally best to SELECT only the most specific parts of the name and ignore the common words with no subject content. Abbreviations may be used. For example, to find documents issued by the various divisions of the Minnesota State Dept. of Education,

(¶ 1.16)

```
? SELECT MINNESOTA(W)STATE(2W)EDUCATION/CS
          5    154 MINNESOTA(W)STATE(2W)EDUCATION
? TYPE 5
1
ED134491  SO009706
  Political Involvement. Community *Involvement/Career Education: An
Experience-based Social Studies Program, Volume III.
  Kyle, James; And Others
  Minnesota State Dept. of Education, St. Paul. Div. of Vocational and
Technical Education.; Robbinsdale Independent School District 281,
Minn.
  Publ. Date: Jan 75 Note: 162p.; For related document, see SO 009
705; Many pages of the original document are copyrighted and therefore
not available. They are not included in the pagination
  EDRS Price MF-$0.83 HC-$8.69 Plus Postage.
  Descriptors: *Career Education/ Career Planning/ Community
Involvement/ Democratic Values/ Educational Objectives/ Individual
Instruction/ Instructional Materials/ Learning Activities/ Political
Power/ *Politics/ Relevance (Education)/ Secondary Education/ *Social
Studies/ Social Studies Units/ Student Participation/ Student
Volunteers/ Study Guides
```

DESCRIPTOR (/DE, /DE*, /DF, /DF*). Descriptors are single-word or multiword subject terms assigned from the *Thesaurus of ERIC Descriptors* to characterize the substantive content of the original document. On the average, ten descriptors are assigned to each indexed document in the RIE portion of the database, and fewer descriptors are assigned to the articles in the CIJE portion of the database. The assignment of descriptors from the *Thesaurus* is discussed in detail in the *ERIC Processing Manual.* The *Thesaurus* is reissued annually to incorporate vocabulary changes.

(¶ 1.17)

Thesaurus terms and the relationships among them have been incorporated into the online index as an aid to DIALOG system searchers. It is possible to view the relationships by further EXPANDing any term in the basic index that shows postings under the "RT" heading, as illustrated below.

(¶ 1.18)

```
? EXPAND REMEDIAL READING
Ref   Index-term                Type  Items  RT
E1    REMEDIAL EDUCATION
        PROGRAMS----------                   1
E2    REMEDIAL HANDWRITING----               1
E3    REMEDIAL INSTRUCTION----             712  16
E4    REMEDIAL MATHEMATICS----             164   6
E5    REMEDIAL PROGRAMS-------             586  11
E6   -REMEDIAL READING-------             889   8
E7    REMEDIAL HEADING CLINICS              38   3
E8    REMEDIAL READING PROGRAM               1
E9    REMEDIAL READING
        PROGRAMS----------                 369   4
E10   REMEDIAL STRATEGIES-----              1
E11   REMEDIAL TEACHERS-------             41   2
E12   REMEDIAL TEACHING-------              1
E13   REMEDIAL WRITING--------              1
E14   REMEDIALLY--------------              2
E15   REMEDIALS---------------              3
E16   REMEDIALSUMMER----------              1
E17   REMEDIATE---------------             62
E18   REMEDIATED--------------              8
                                         -more-
? EXPAND E6
Ref   Index-term                Type  Items  RT
R1   -REMEDIAL READING--------            889   8
R2    CORRECTIVE READING------N             55   1
R3    READING-----------------B          22699  71
R4    REMEDIAL COURSES--------B             84   6
R5    READING CENTERS---------R            176   4
R6    REMEDIAL INSTRUCTION----R            712  16
R7    REMEDIAL READING CLINICSR             38   3
R8    REMEDIAL READING
        PROGRAMS----------R                369   4
R9    RETARDED READERS--------R            227   5
```

The following term relationships are coded in the "Type" column:

(¶ 1.19)

B - Broader term

N - Narrower term

R - Related term

U - Use or Used for

FILE 1

Terms such as E1 above (without postings in the basic index but with a related-term posting) are included to provide cross references to proper *Thesaurus* terms. Thus, the searcher should EXPAND the E reference number to locate the preferred term, e.g.,

(¶ 1.20)

```
? EXPAND E1
Ref   Index-term              Type Items RT
R1   -REMEDIAL EDUCATION
       PROGRAMS----------                1
R2    REMEDIAL PROGRAMS-------U    586 11
```

Here the "U" indicates that the searcher should use "remedial programs" rather than "remedial education programs."

(¶ 1.21)

If the searcher is familiar with the *Thesaurus*, the related terms can be EXPANDed directly by enclosing the term in parentheses, e.g.,

(¶ 1.22)

```
? EXPAND (LITERATURE APPRECIATION)
Ref   Index-term              Type Items RT
R1   -LITERATURE APPRECIATION-    1093  4
R2    READING ENJOYMENT          U      1
R3    CHORAL SPEAKING----------R     44  7
R4    LITERATURE---------------R 16515118
R5    LITERATURE GUIDES-------R    158  4
```

Here the "U" indicates that "literature appreciation" is used for or preferred over "reading enjoyment."

(¶ 1.23)

Some *Thesaurus* terms may include a scope note for further definition. Scope notes are not entered when the term is SELECTed, but appear in the system response, e.g.,

(¶ 1.24)

```
? SELECT READINESS
          6  2345 READINESS (PREPAREDNESS TO RE
```

Long terms or terms with long scope notes are truncated at 42 characters as is the entry above.

(¶ 1.25)

Also, some *Thesaurus* terms may include a parenthetical word or phrase as part of the term. In such a case the parenthetical expression must be included in the entry to restrict retrieval to that specific term, e.g.,

(¶ 1.26)

```
? SELECT READINESS (MENTAL)
          7    122 READINESS (MENTAL)
```

(However, the citations in set 7 would also be included in the single-word index entry in set 6.)

(¶ 1.27)

As many as six descriptors may appear with an asterisk, indicating that they are the "major" descriptors. (The major descriptors are used as subject index entries for the printed issues of RIE and CIJE.) Searching can be restricted to the major descriptors by the use of the /DE* and /DF* suffixes as illustrated below or by the use of the LIMIT command discussed on page 1-20.

(¶ 1.28)

When the searchable indexes are created, all multiword descriptors are re-indexed under each meaningful individual word of the phrase. In addition, it is possible to differentiate between single-word descriptors and the same single words extracted from multiword descriptors. The following example with the term "reading" illustrates the varying degrees of specificity and differences in postings.

(¶ 1.29)

```
? SELECT READING; SELECT READING/DE; SELECT READING/DE*
          8 22699 READING
          9 17104 READING/DE
         10 12555 READING/DE*
? SELECT READING/DF; SELECT READING/DF*
         11  2538 READING/DF
         12   864 READING/DF*
```

In set 8, the term is retrieved from any suffix-coded field in the basic index. In set 9, the term is retrieved from single-word or multiword descriptors, such as "reading," "reading ability," "reading achievement," "reading instruction," "critical reading," "supplementary reading materials," etc. However, in set 10 the term is retrieved from these descriptors only if they have been asterisked as major. In sets 11 and 12 retrieval is restricted to the single-word term READING, and in the latter instance the word must have been asterisked as a major descriptor.

(¶ 1.30)

When the multiword descriptors are re-indexed under the individual words of the phrase, a few common words without subject content are omitted from the index. To search for phrases containing such "stop" words, the descriptor phrase can be SELECTed as a unit, e.g.,

(¶ 1.31)

```
? SELECT TRANSFER OF TRAINING
         13   802 TRANSFER OF TRAINING
```

However, if full-text methods are used, the "of" must be excluded, but counted, e.g.,

(¶ 1.32)

```
? SELECT TRANSFER(1W)TRAINING/DE
         14   802 TRANSFER(1W)TRAINING/DE
```

The search can be broadened by removing the restriction to the Descriptor field, e.g.,

(¶ 1.33)

```
? SELECT TRANSFER(1W)TRAINING
         15   833 TRANSFER(1W)TRAINING
```

The additional records retrieved by set 15 have the phrase in the Abstract, Identifier, or Title fields.

(¶ 1.34)

FILE 1

IDENTIFIER(/ID, /ID*, /IF, /IF*). Identifiers are single-word or multiword terms assigned to provide additional subject indexing beyond descriptor terms. Identifiers are generally free-language terms and tend to be more specific than the *Thesaurus* vocabulary, e.g., project names, legislation, geographic locations, and other subject terms which have not yet achieved broad acceptance. As many as two identifiers per record may be asterisked to designate them as "major."

(¶ 1.35)

When identifier phrases are entered into the basic index, they are allowed a maximum of 42 characters in length. To search for lengthy phrases, the searcher can utilize the truncation capability, e. g.,

(¶ 1.36)

```
? SELECT READING AND MATHEMATICS OBSERVATION SYS?
          16      4 READING AND MATHEMATICS OBSERV
? TYPE 16
1
EJ149053  SP505053
  Teacher Aptitudes, Knowledge, Attitudes, and Cognitive Style as
Predictors of Teaching Behavior
  Ekstrom, Ruth B.     Journal of Teacher Education; 27; 4; 329-31
Win 76
  Descriptors: *Teacher Behavior/ *Teacher Attitudes/ *Teaching Styles
/ *Teacher Characteristics/ *Elementary School Teachers/ *Cognitive
Ability/ Teacher Qualifications/ Predictor Variables/ Aptitude/ Verbal
Ability
  Identifiers: *Beginning Teacher Evaluation Study/ APPLE Observation
System/ Reading and Mathematics Observation System (RAMOS)
```

However, since identifiers are also indexed under each meaningful individual word, full-text methods can be used, e.g.,

(¶ 1.37)

```
? SELECT READING(1W)MATHEMATICS(W)OBSERVATION/ID
          17      4 READING(1W)MATHEMATICS(W)OBSER
```

(Note that while the command has been processed properly, the resulting set title is shortened.)

(¶ 1.38)

Geopolitical identifiers currently begin with the name of a state (U.S.), province (Canada), or country (all others) and include the local name as a parenthetical qualifier.

(¶ 1.39)

However, prior to January 1976, the area and local names were entered as separate identifiers. Thus, if a search involves a specific location, both formats can be retrieved by use of the F full-text operator, e.g.,

(¶ 1.40)

```
? SELECT MEMPHIS(F)TENNESSEE/ID
          18     23 MEMPHIS(F)TENNESSEE/ID
? TYPE 18
1
EJ148118  EC090028
  The Educational, Behavioral, and Psychological Characteristics of
Teachers as Motivators for Creative and Less Creative Students
```

Hooks, Mose Yvonne Brooks North Carolina Association for the Gifted and Talented Quarterly Journal; 1; 4; 52-60 F 75
Descriptors: *General Education/ *Creative Ability/ *Creativity Research/ *Teacher Characteristics/ *Student Motivation/ Research Projects/ Elementary Secondary Education
Identifiers: Tennessee (Memphis)

? TYPE 18/2/23
23
ED001245
THE MEMPHIS STORY-MOBILIZING A LARGE SCHOOL SYSTEM AND A COMMUNITY FOR THE SUPERIOR AND TALENTED STUDENT PROJECT.
COUNCE, D. SHELBY
North Central Association of Coll. and Secondary Schools, Chicago, Ill.; Southern Association of Colleges and Schools, Atlanta, Ga.
Publ. Data: 65 Note: 68P.
EDRS PRICE MF-$0.76 HC-$3.32 PLUS POSTAGE
Descriptors: Academic Aspiration/ Curriculum Development/ *Guidance Services/ *Identification/ Occupational Choice/ Special Programs/ *Student Motivation/ *Superior Students/ *Talented Students
Identifiers: ATLANTA GEORGIA/ CHICAGO/ ILLINOIS/ MEMPHIS/ STS PROJECT/ TENNESSEE

In actual searching, it would be best to include the other relevant suffix-coded fields in the entry, e.g.,

(¶ 1.41)

```
? SELECT MEMPHIS(F)TENNESSEE/ID,AB,TI
          19     65 MEMPHIS(F)TENNESSEE/ID,AB,TI
```

The above entry is generally better for geographic area searching than an unsuffixed entry which would retrieve also from the /NT, /CS, and /SA fields.

(¶ 1.42)

Personal names may be used as identifiers and generally are entered with surname first. However, for comprehensive searching, the entered term should have a format compatible with all fields in which it is likely to occur. For very well-known names, the surname may be specific enough, e.g., SELECT SHAKESPEARE. For less unique surnames, the initials and/or name should be included in the entry, e.g., SELECT E(W)B(F)WHITE.

(¶ 1.43)

When the searchable indexes are created, each identifier are indexed as the whole phrase and re-indexed under each meaningful individual word of the phrase. In order to differentiate between single-word identifiers and the same single words extracted from multiword identifiers, the /IF and /IF* suffixes can be used. For example,

(¶ 1.44)

```
? SELECT STANDARDIZATION/ID
          20     13 STANDARDIZATION/ID
```

retrieves such identifiers as "standardization," "test standardization," and "international organization for standardization," while

(¶ 1.45)

FILE 1

```
? SELECT STANDARDIZATION/ID*
          21      5 STANDARDIZATION/ID*
```

retrieves these same identifiers only if they have been asterisked as major, but

(¶ 1.46)

```
? SELECT STANDARDIZATION/IF
          22      7 STANDARDIZATION/IF
```

retrieves only the single-word identifier, and

(¶ 1.47)

```
? SELECT STANDARDIZATION/IF*
          23      2 STANDARDIZATION/IF*
```

retrieves the single-word identifier only if it has been asterisked as major.

(¶ 1.48)

DESCRIPTIVE NOTE (/NT). This field contains pagination and other cataloging information to augment the document description, such as the circumstances under which the document was prepared (conference paper, speech, reprint, preprint, etc.), the language of the document, relation to other documents, and legibility of the document. This field is not used for CIJE documents.

(¶ 1.49)

SPONSORING AGENCY (/SA). The Sponsoring Agency field is used to list the institution, e.g., government agency, private foundation, etc., other than the corporate author, which supported the work or production of the document by providing funds via contract or grant. This field is used only for RIE items. Each sponsoring agency name has been entered in a standardized form and corresponds to a specific eight-character alphanumeric code searchable in the SC= field. When the sponsoring agency is also the corporate author, this field is not used.

(¶ 1.50)

In searching the /SA field, full-text methods should be used with the most unique words of the name. For example,

(¶ 1.51)

```
? SELECT CHILD(W)DEVELOPMENT/SA
          24     452 CHILD(W)DEVELOPMENT/SA
```

retrieves the records of reports sponsored or cosponsored by the Office of Child Development (DHEW), Washington, D.C.

(¶ 1.52)

Since the Sponsoring Agency field is not used if the report is sponsored by the issuing agency, the searcher is likely to want to retrieve reports sponsored or issued by a particular agency. This is best achieved by using both suffixes with the entry, e.g.,

(¶ 1.53)

```
? SELECT CHILD(W)DEVELOPMENT/CS,SA
      25    700 CHILD(W)DEVELOPMENT/CS,SA
```

TITLE (/TI). The complete title is entered as it appears on the original document, including alternative title, subtitle, and other associated descriptive matter. Fabricated titles, for documents without an original title, are enclosed in brackets.

(¶ 1.54)

All meaningful individual words from the Title are searchable. Any hyphenated words are separated into their component parts and must be searched using full-text methods. Thus, to find titles with the term "non-English speaking," the following command could be entered.

(¶ 1.55)

```
? SELECT NON(W)ENGLISH(W)SPEAKING/TI
         26     46 NON(W)ENGLISH(W)SPEAKING/TI
? TYPE 26/6/1-2
1
EJ151279
  The Non-English Speaking Child in Your Classroom

2
EJ146312
  The Puerto Rican Non-English-Speaking Child: What Can I Do?
```

Note that both the hyphenated and non-hyphenated forms are retrieved.

(¶ 1.56)

LEGISLATIVE AUTHORITY CODE (AC=). For RIE documents resulting from research funded by the Office of Education or by the National Institute of Education, a two-digit code indicates the legislative authority under which the project was funded. The following codes are used in this field:

(¶ 1.57)

CODE	PROGRAM AREA
08	Adult and Vocational Education, PL 88-210
16	Captioned Films for the Deaf, PL 85-905
24	Cooperative Research, PL 89-10, Title IV
32	Disadvantaged Students Program, PL 89-10, Title I
40	Handicapped Children and Youth, PL 88-164
48	Language Development, PL 85-864, Title VI
52	Library Research and Development, PL 89-320, Title III, Part B
56	New Educational Media, PL 85-864, Title VII, Part A
64	New Educational Media, PL 85-864, Title VII, Part B

FILE 1

CODE	PROGRAM AREA
72	Research in Foreign Countries, PL 83-480
80	State Educational Agencies Experiment Activities, PL 89-10, Title V, Section 505
88	Supplementary Centers and Services, PL 88-10, Title III
95	Other Office of Education and National Institute of Education Programs (Projects for which the legislative authority cannot be determined are included here.)

The legislative authority code appears after the ED number, as illustrated in the following example.

(¶ 1.58)

```
? SELECT AC=40
          27    338 AC=40
? TYPE 27/2/2
2
ED129532  40  RC009498
  Outdoor Education for Emotionally Handicapped Students in Central
New York. Final Report, September 1975. A Report on Project 75-940 for
Handicapped Children Under Title VI-B (August 1, 1974-Jul 31, 1975).
  Miller, James
  Publ. Date: Sep 75  Note: 70p.
  EDRS Price MF-$0.83 HC-$3.50 Plus Postage.
  Descriptors: Communication Skills/ Comparative Analysis/ Educational
Objectives/ *Effective Teaching/ *Emotionally Disturbed/ Environmental
Education/ Formative Evaluation/ Motor Development/ *Outdoor Education
/ Participant Involvement/ Pilot Projects/ *Program Evaluation/ Self
Concept/ Social Development/ Special Education/ *Student Attitudes/
*Student Behavior/ Student Development/ Teacher Evaluation
  Identifiers: Elementary Secondary Education Act Title VI B/ ESEA
Title VI B/ *New York (Central)
```

AUTHOR (AU=). Up to two individual authors per document may be listed in an online record. If there are more than two authors, only the first author is listed followed by "And Others." Author names generally appear as published in the source document. The AU= index has been EXPANDed to illustrate the format of the entries.

(¶ 1.59)

```
? EXPAND AU=MOORE, JOHN W.
Ref   Index-term              Type  Items RT
E1    AU=MOORE, JOHN M.-------      2
E2    AU=MOORE, JOHN N.-------      3
E3    AU=MOORE, JOHN P.-------      1
E4    AU=MOORE, JOHN R.-------      1
E5    AU=MOORE, JOHN ROBERT---      1
E6    AU=MOORE, JOHN W.-------      8
E7    AU=MOORE, JOHN W., ED.--      7
```

```
E8    AU=MOORE, JOHNNA--------        1
E9    AU=MOORE, JOSEPH A.-----        3
E10   AU=MOORE, JOSEPH H.-----        5
E11   AU=MOORE, JOSEPH T.-----        1
E12   AU=MOORE, JOSIAH--------        1
E13   AU=MOORE, JULIE L.------        2
E14   AU=MOORE, JUNE----------        6
E15   AU=MOORE, JUSTIN R.-----        1
E16   AU=MOORE, K. E.---------        1
E17   AU=MOORE, K. L.---------        1
E18   AU=MOORE, KATE----------        1
E19   AU=MOORE, KATHRYN M.----        4
                                  -more-
```

Author names may include full name or initials, or some combination of name(s) and initials with spacing and punctuation. Variations exist, depending on the source document. If the individual edited, rather than wrote, the document, "ED." is appended to the name as in the entry at E7. EXPANDing is recommended prior to SELECTing an author name.

(¶ 1.60)

```
? SELECT E6, E7
          28    15 E6, E7
                   E6: AU=MOORE, JOHN W.
```

CLEARINGHOUSE CODE (CH=). Currently there are 16 clearinghouses located at various universities and professional organizations throughout the United States which are responsible for acquiring, indexing and abstracting the reports and journal articles relevant to their respective subject area specialty.

(¶ 1.61)

Each clearinghouse assigns a temporary accession number to each document it inputs. The clearinghouse accession numbers searchable in the CH= field consist of a two-letter clearinghouse code followed by a six-digit number, e.g., CH=SE506669.

(¶ 1.62)

All clearinghouse accession numbers have been cascaded to the two-letter clearinghouse code, e.g., SELECT CH=SE retrieves all documents indexed by the Science, Mathematics, and Environmental Education Clearinghouse. The clearinghouse code can be used as a broad subject area restriction in a search strategy as illustrated below.

(¶ 1.63)

```
? SELECT SECONDARY EDUCATION
          29  1293 SECONDARY EDUCATION   (EDUCATIO
? SELECT SECONDARY(W)EDUCATION(C)CH=SE
          30  3153 SECONDARY(W)EDUCATION(C)CH=SE
```

In set 30 the C full-text operator has been used to restrict the term "secondary education" to documents indexed by the Science, Mathematics, and Environmental Education Clearinghouse.

(¶ 1.64)

Over the timespan of the database, there have been some changes in the number of clearinghouses as well as in their codes. The current clearinghouse codes and previous codes are listed in Table 1. To search in a

subject area where the clearinghouse was discontinued and merged with another, SELECT the old code and the current code. For example, for a broad search on Library and Information Sciences.

(c 1.65)

```
? SELECT CH=LI, SELECT CH=IR, COMBINE 31 OR 32
          31   6228 CH=LI
          32   8506 CH=IR
          33  14734   31 OR 32
```

TABLE 1. Clearinghouse Codes (CH=)

AA - ERIC Facility Contractor (series discontinued with March 1973 RIE)

AC - Adult Education (merged into CE in 1973)

AL - Linguistics (merged into FL in 1971)

CE - Career Education

CG - Counseling and Personnel Services

CS - Reading and Communication Skills

EA - Educational Management

EC - Handicapped and Gifted Children

EF - Educational Facilities (merged into EA in 1970)

EM - Educational Media and Technology (merged into IR In 1974)

FL - Languages and Linguistics

HE - Higher Education

IR - Information Resources

JC - Junior Colleges

LI - Library and Information Sciences (merged into IR in 1974)

PS - Early Childhood Education

RC - Rural Education and Small Schools

RE - Reading (merged into CS in 1972)

SE - Science, Mathematics, and Environmental Education

SO - Social Science/Social Studies

SP - Teacher Education

TE - Teaching of English (merged into CS in 1972)

TM - Tests, Measurement, and Evaluation

UD - Urban Education

VT - Vocational and Technical Education (merged into CE in 1973)

CONTRACT/GRANT NUMBER (CN=). If the contract or grant numbers assigned by sponsoring agencies are plainly apparent on the document, they are included in this field. Prior to July 1976, only U.S. Office of Education (OE) or National Institute of Education (NIE) grant and contract numbers were indexed in this field.

(¢ 1.66)

Contract and grant numbers generally include an alphabetic prefix followed by a variable number of digits. Some typical contract and grant numbers are:

(¢ 1.67)

OE-0-73-7094

OCD-CB-02(C4)

NIE-C-400-75-0015

N61339-73-C-0097

All contract and grant numbers are cascaded to the initial alphabetic prefix. Thus SELECT CN=OCD includes CN=OCD-CB-02(C4) and all other contract numbers beginning with "OCD."

(¢ 1.68)

This field is not used for CIJE items.

(¢ 1.69)

DOCUMENT TYPE (DT=). Since September 1974, a single-character alphabetic code has been used in this field to indicate the form of the RIE document. The codes and their corresponding publication types are listed in Table 2.

(¢ 1.70)

SELECTion of a document type code automatically restricts retrievel to only those RIE documents entered in the ERIC database since September 1974.

(¢ 1.71)

To find recent review articles or bibliographies on the "gifted," the following command could be entered.

(¢ 1.72)

```
? SELECT DT=L(C)GIFTED
          34     30 DT=L(C)GIFTED
```

TABLE 2. Document Type Codes (DT=)

A	Audio Visual/Nonprint Media; Audiovisual Aids; Films; Tape Recordings; Phonotape Recordings; Computer Programs; etc.
B	Books; Monographs, Textbooks; Programmed Texts; etc. (not otherwise classifiable)
C	Curriculum Guides; Teacher-Developed Materials; Laboratory Manuals
D	Directories; Membership Lists; Tables of Organization; Reference Works Dealing with Organizations/Institutions; etc.
G	Guides; Teaching Guides; Resource Guides; Study Guides; Administrative Guides; Leaders Guides; Manuals; Training Manuals
H	Legislation, Legislative Hearings, Legislative Reports, Congressional Documents. (Include both Federal and State levels; include National Commissions).
J	Journal Articles; Serials; Periodicals; Bulletins; Newsletters; Newspapers; etc.
K	Program/Project Descriptions
L	Bibliographies; Annotated Bibliographies; Book Catalogs; Abstracts; Literature Reviews; Literature Searches/Guides; Book Lists; Book Reviews; Library Guides; Indexes (Locators); State-of-the-Art Reviews
M	Maps; Atlases; Gazetteers
N	Numerical and Statistical Tables; Quantitative Data and Analyses
O	Other
P	Proceedings; Conference Records/Minutes (entire)
Q	Questionnaires; Tests; Measurement Devices; Evaluation Devices
R	Reports (Research and Technical)
S	Speeches; Conference Reports; "Papers presented at. . .", Verbal Presentations
T	Theses; Dissertations
V	Dictionaries; Vocabularies; Glossaries; Thesauri
Y	Annual Reports; Yearbooks

GROUP CODE (GC=). Specific *Thesaurus* terms appropriate to very broad subject areas have been grouped together in the Group Code field. Each code is a three-digit number and represents a group of related subject terms, e.g., GC=060 for Behavior terms.

(¶ 1.73)

Group codes cannot be SELECTed directly. The EXPAND command can be used with a group code in parentheses to see the terms included in the group and their associated postings. Individual terms, or all terms included in the group may be SELECTed by their R reference numbers as illustrated below.

(¶ 1.74)

```
? EXPAND (GC=360)
Ref   Index-term              Type  Items  RT
R1   -GC=360  (OPPORTUNITIES)-          14
R2    AFFIRMATIVE ACTION------D    293  22
R3    CAREER OPPORTUNITIES----D   1037  10
R4    COMMUNITY BENEFITS------D     87   7
R5    CULTURAL OPPORTUNITIES--D     55   5
R6    ECONOMIC OPPORTUNITIES--D    165   3
R7    EDUCATIONAL
        OPPORTUNITIES---------D   1770   9
R8    EMPLOYMENT OPPORTUNITIESD  2450  29
R9    HOUSING OPPORTUNITIES---D     69   4
R10   OPPORTUNITIES-----------D 10742  10
R11   RESEARCH APPRENTICESHIPSD    23   2
R12   RESEARCH OPPORTUNITIES--D    254   2
R13   SOCIAL OPPORTUNITIES----D     80   8
R14   TEACHING BENEFITS-------D     58   2
R15   YOUTH OPPORTUNITIES-----D    160   4
? SELECT R1-R15
          35 11035 R1-R15
                   R1: GC=360 (OPPORTUNITIES)
```

The 52 broad descriptor groups and their codes are listed with their respective scope notes in the *Thesaurus of ERIC Descriptors*.

(¶ 1.75)

ISSUE (IS=). This field is used to indicate the monthly issue of RIE in which the document is announced, e.g., IS=RIEAPR77 for all documents announced in the April 1977 issue of RIE. For documents announced in CIJE, only the year of announcement is indicated, e.g., IS=CIJE1977.

(¶ 1.76)

JOURNAL NAME (JO=). Currently, full journal titles, with the exception of initial articles, are entered in this field. Until 1971, periodical titles were abbreviated according to ANSI Standard Z39.5-1963 "Periodical Title Abbreviations." For comprehensive searching on a specific journal title, both forms should be SELECTed, e.g.,

(¶ 1.77)

```
? SELECT JO=PUBLIC OPIN QUART; SELECT JO=PUBLIC OPINION QUARTERLY
          36     24 JO=PUBLIC OPIN QUART
          37     42 JO=PUBLIC OPINION QUARTERLY
? COMBINE 36 or 37
          38     66  36 OR 37
```

This field may be used for RIE documents.

(¶ 1.78)

BUREAU/PROJECT NUMBER (PN=). This field is used to list the alphanumeric code assigned by the sponsoring agency to the project for which the document was produced. Currently only OE or NIE project numbers are indexed.

(¶ 1.79)

Project numbers generally include an alphabetic prefix followed by a variable number of digits and/or additional letters. Some typical project numbers are:

(¶ 1.80)

BR-H-127245B

CG8119A/5

L0008JA

V-361057L

All project numbers are cascaded to the initial alphabetic prefix. Thus SELECT PN=L includes PN=L0008JA and all other project numbers beginning with L.

(¶ 1.81)

This field is not used for CIJE items.

(¶ 1.82)

REPORT NUMBER (RN=). This field is used to index the unique identifying number(s) assigned to the publication by the organization(s) producing or issuing the document.

(¶ 1.83)

Report numbers generally include an alphabetic prefix followed by a variable number of digits and/or additional letters. Some typical report numbers are:

(¶ 1.84)

AAAS-MISC-PUB-76-2

DA-PAM-550-152

RX74-15-HEW

T-75-196-G

All report numbers are cascaded to the initial alphabetic prefix. Thus SELECT RN=RX includes RN=RX74-15-HEW and all other report numbers beginning with RX. Some entry variations in format of report numbers may exist, so it is best to EXPAND in the RN= index to locate a specific report by its number.

(¶ 1.85)

This field is not used for CIJE items.

(¶ 1.86)

CORPORATE SOURCE CODE or SPONSORING AGENCY CODE (SC=). An eight-character alphanumeric code assigned by ERIC to each originating institution and each sponsoring agency. An alphabetic directory of the institutions, sponsoring agencies and their codes is available (see the Search Aids section). Codes include three letters followed by five digits, e.g., SC=AHP00825 for the Alabama State Dept. of Education, Montgomery, or SC=BBB08152 for the Office of Education (DHEW), Washington, D.C., Right to Read Program.

(¶ 1.87)

This field is not used for CIJE documents.

(¶ 1.88)

UPDATE (UD=). Search results can be restricted to a given update of the database, e.g., SELECT UD=7705. In addition to the individual update numbers, UD=9999 is always identical to the latest update in the database. A SELECTed update set can be COMBINEd with a subject set to retrieve citations in that update of ERIC.

(¶ 1.89)

YEAR (YR=). This field is used to indicate the last two digits of the year the document was announced in RIE or CIJE. However, to restrict retrieval to a range of years, it is more convenient to use the LIMIT command as described in the following section.

(¶ 1.90)

LIMITING

The LIMIT command is used to reduce a set according to some criterion applicable to the database. Generally this command should be entered after SELECTing the appropriate set. Sets SELECTed in the ERIC database may be LIMITed by subfile, by accession number within each subfile, by document availability, and by major or minor descriptor or identifier.

(¶ 1.91)

ED or EJ Subfile. A SELECTed set can be restricted to the report literature subfile (RIE) or the periodical literature subfile (CIJE) by use of the /ED or /EJ suffixes respectively as illustrated below.

(¶ 1.92)

```
? SELECT SCIENCE(W)EDUCATION
          39 11003 SCIENCE(W)EDUCATION
? LIMIT 39/ED; LIMIT 39/EJ
          40  3649  39/ED
          41  7354  39/EJ
```

Accession Number. The DIALOG accession number which appears in each printed record is an eight-character number which includes the ED or EJ prefix followed by six digits and corresponds to the accession number which appears on each record in the printed RIE or CIJE. To LIMIT by accession number range, the prefix is dropped and the appropriate suffix is used with the accession number range.

(¶ 1.93)

The annual accession number ranges that may be used with the LIMIT command for each subfile are listed below. This table is also available online in response to ?YEARS1.

(¶ 1.94)

YEAR	ED NUMBERS	EJ NUMBERS
PRE-66	002747-003960/ED	
1966	010000-010093/ED	
1967	010094-012348/ED	
1968	012349-021151/ED	
1969	021152-031604/ED	000001-011707/EJ
1970	031605-042060/ED	011708-027599/EJ
1971	042061-054390/ED	027600-045271/EJ
1972	054391-066620/ED	045272-062751/EJ
1973	066621-080787/ED	062752-082164/EJ
1974	080788-095253/ED	082165-101872/EJ

FILE 1

1975	095254-110594/ED	101873-121926/EJ
1976	110595-127413/ED	121927-142252/EJ
1977	127414-	142253-

To restrict a set to a given time period, the appropriate DIALOG accession number range for each subfile is used and the resulting sets are COMBINEd using the Boolean OR. For example, to LIMIT set 39 to those documents announced in ERIC from 1974 to the present:

(¶ 1.95)

```
? LIMIT 39/080788-999999/ED; LIMIT 39/082165-999999/EJ
          42   2335   39/080788-999999/ED
          43   5576   39/082165-999999/EJ
? COMBINE 42 OR 43
          44   7911   42 OR 43
```

In LIMITing by accession number range, it is necessary to have all of the digits filled with numbers. To retrieve citations including the very newest entries, it is possible to "fill" the upper range with nines as shown in the previous example.

(¶ 1.96)

Document Availability. A SELECTed set can be restricted to those documents for which full copies are available or not available from the ERIC Document Reproduction Service (EDRS) by the use of the suffixs /AVAIL and /UNAVAIL respectively.

(¶ 1.97)

```
? LIMIT 39/AVAIL; LIMIT 39/UNAVAIL
          45   2586   39/AVAIL
          46   8417   39/UNAVAIL
```

Since copies of documents announced in the CIJE subfile are not available from EDRS, the available documents are always from the RIE (or ED) subfile.

(¶ 1.98)

Major or Minor Descriptor or Identifier. The asterisked descriptors and identifiers (which are also the subject index entries in the printed RIE and CIJE) are considered to be major terms. A search may be LIMITed to major terms only by using /MAJ. The /MIN suffix LIMITs the set to all other occurrences of the term.

(¶ 1.99)

The /MAJ and /MIN LIMIT suffixes are properly applied to sets formed from terms SELECTed in the Descriptor and/or Identifier fields.

(¶ 1.100)

However, it is generally preferable to restrict a search to major descriptor or identifier by using the /DE* and /ID* fields, e.g.,

(¶ 1.101)

```
? SELECT SCIENCE(W)EDUCATION/DE*,ID*
          47   2825 SCIENCE(W)EDUCATION/DE*,ID*
```

Multiple Limits. Multiple LIMITs can be used if they are entered in the acceptable order, namely the accession number range precedes the subfile, which precedes the major/minor term, which precedes the

availability/nonavailability specification. For example, to apply all of the LIMIT restrictions to the set on Science Education, the following entry can be used:

(¶ 1.102)

```
? LIMIT 39/080788-999999/ED/MAJ/AVAIL
          48    480  39/080788-999999/ED/MAJ/AVAIL
```

However, the above entry is redundant. Since the availability LIMIT restricts the set to only ED documents it is not necessary to include the ED suffix in the entered command, e.g.,

(¶ 1.103)

```
? LIMIT 39/080788-999999/MAJ/AVAIL
          49    480  39/080788-999999/MAJ/AVAIL
```

The general order listed above must be followed even when only some of the LIMIT specifications are used, e.g.,

(¶ 1.104)

```
? LIMIT 39/080788-999999/ED/MAJ
          50    647  39/080788-999999/ED/MAJ
? LIMIT 39/MAJ/AVAIL
          51    839  39/MAJ/AVAIL
```

LIMITALL. The LIMITALL command can be used in two ways: to LIMIT to an accession number range or to restrict the search to ED or EJ documents. The LIMITALL command can be used with the /ED or /EJ suffix to restrict all subsequent sets to the subfile specified by the suffix, until the command is canceled with the LIMITALL/ALL command, or a new LIMITALL command, e.g.,

(¶ 1.105)

```
? LIMITALL/ED
          LIMIT ALL ALL/ED
? SELECT SCIENCE(W)EDUCATION
          52   3649 SCIENCE(W)EDUCATION
```

The accession number range can be entered with the LIMITALL command to restrict all subsequent sets to the time period specified by the accession numbers within the appropriate subfile, until the command is canceled with the LIMITALL/ALL command, or a new LIMITALL command, e.g.,

(¶ 1.106)

```
? LIMITALL/080788-999999/ED
          LIMIT ALL ALL/080788-999999/ED
? SELECT SCIENCE(W)EDUCATION
          53   2335 SCIENCE(W)EDUCATION
? LIMITALL/ALL
          LIMIT ALL ALL/ALL
```

FILE 1

FORMAT OPTIONS

If a format is not specified in TYPE, DISPLAY, or PRINT commands, format 2 is assumed. The records are displayed with the highest accession number (newest record) first, within each subfile of File 1. All CIJE (EJ) records precede all RIE (ED) records.

(¶ 1.107)

FORMAT 1: DIALOG Accession Number

(¶ 1.108)

```
EJ034210  EJ025800  EJ021833  EJ013689  EJ009626  EJ008710  ED132588
ED131621  ED131416  ED127236  ED126505  ED099878  ED099001  ED098968
```

FORMAT 2: Complete Record Except Abstract

(¶ 1.109)

CIJE

```
EJ150129  CS710249
  The Cloze Procedure: Can It Be Used with Literature?
  Dupuis, Mary M.    Reading Improvement; 13; 4; 199-203    W 76
  Descriptors: *Cloze Procedure/ *Readability/ *Literature/ *Reading
Comprehension/  *Reading  Ability/  *Predictive  Validity/  Reading
Research/ Secondary Education/ Grade 10
```

RIE

```
ED131416=  CS003014
  Reading Guidance in a Media Age.
  Polette, Nancy; Hamlin, Marjorie
  Publ. Date: 75  Note: 260p.
  Available from: Scarecrow Press, Inc., 52 Liberty St., P.O. Box 656,
Metuchen, New Jersey 08840 ($10.00 cloth)
  Document Not Available from EDRS.
  Descriptors: Audiovisual Aids/ *Childrens Books/ *Childrens
Literature/ *Early Reading/ Elementary Education/ Elementary School
Libraries/ *Literacy/ *Mass Media/ Reading Development/ *Reading
Instruction/ Reading Materials/ Teaching Techniques
```

FORMAT 3: Bibliographic Citation

(¶ 1.110)

CIJE

```
EJ150129
  The Cloze Procedure: Can It Be Used with Literature?
  Dupuis, Mary M.
  Reading Improvement, 13, 4, 199-203
  W 76
```

FILE 1

RIE

ED131416#
Reading Guidance in a Media Age.
Polette, Nancy; Hamlin, Marjorie
260p. 75
Scarecrow Press, Inc., 52 Liberty St., P.O. Box 656, Metuchen, New Jersey 08840 ($10.00 cloth)
Document Not Available from EDRS.

FORMAT 4: Abstract

CIJE (¶ 1.111)

EJ150129
Reveals that the cloze and comprehension test scores correlated at significant levels, thus supporting the predictive validity of the cloze procedure. (RB)

RIE

ED13141#
Chapters in this volume discuss the problems that teachers, parents, and librarians face in encouraging elementary school children to read and suggest ways to improve student motivation and reading ability in this age of mass media. "Sparking the Fuse for Literacy" defines the complex term "literacy," and examines the effects of the media on reading. "Home is Where the Start Is" discusses parental roles in reading development, while "Library Lures to Literacy" suggests activities to stimulate student interest. Other techniques designed to enliven the reading process are provided in "Independent Study Programs in Literature," "From Page to Stage--Via Drama and Puppetry," and "Sharing Literature: The Audiovisual Experience." "Sound, Visuals, Action" explores the development of a student production center using video and audio equipment. Conference planning and introducing authors to children through audiovisual media are discussed in "Students, Teachers, Librarians and Authors Get Together." In addition, each chapter contains a brief bibliography of children's books. (KS)

FORMAT 5: Full Record

CIJE (¶ 1.112)

EJ150129 CS710249
The Cloze Procedure: Can It Be Used with Literature?
Dupuis, Mary M. Reading Improvement; 13; 4; 199-203 W 76
Descriptors: *Cloze Procedure/ *Readability/ *Literature/ *Reading Comprehension/ *Reading Ability/ *Predictive Validity/ Reading Research/ Secondary Education/ Grade 10
Reveals that the cloze and comprehension test scores correlated at significant levels, thus supporting the predictive validity of the cloze procedure. (RB)

FILE 1

RIE

ED131416= CS003014
Reading Guidance in a Media Age
Polette, Nancy, Hamlin, Marjorie
Publ Date 75 Note 260p
Available from Scarecrow Press Inc., 52 Liberty St., P O Box 656 Metuchen New Jersey 08840 ($10 00 - oth).
Document Not Available from EDRS
Descriptors Audiovisual Aids *Childrens Books *Childrens Literature *Early Reading Elementary Education Elementary School Libraries *Literacy *Mass Media Reading Development *Reading Instruction Reading Materials Teaching Techniques
Chapters in this volume discuss the problems that teachers, parents, and librarians face in encouraging elementary school children to read and suggest ways to improve student motivation and reading ability in this age of mass media. "Sparking the Fuse for Literacy" defines the complex term "literacy," and examines the effects of the media on reading. "Home Is Where the Start Is" discusses parental roles in reading development, while "Library Lures to Literacy" suggests activities to stimulate student interest. Other techniques designed to enliven the reading process are provided in "Independent Study Programs in Literature," "From Page to Stage--Via Drama and Puppetry," and "Sharing Literature The Audiovisual Experience." "Sound, Visuals Action" explores the development of a student production center using video and audio equipment. Conference planning and introducing authors to children through audiovisual media are discussed in "Students, Teachers, Librarians and Authors Get Together." In addition, each chapter contains a brief bibliography of children's books. (KS)

FORMAT 6: Title

(¶ 1.113)

CIJE

EJ150129
The Cloze Procedure: Can It Be Used with Literature?

CIE

ED131416
Reading Guidance in a Media Age.

Direct Access. Any record in the file may be TYPEd or DISPLAYed directly in any format using the DIALOG Accession Number/Format in the TYPE or DISPLAY command, e.g.,

(¶ 1.114)

? TYPE EJ150129/4

EJ150129
Reveals that the cloze and comprehension test scores correlated at significant levels; thus supporting the predictive validity of the cloze procedure. (RB)

FILE 1

SEARCH AIDS

1. *Thesaurus of ERIC Descriptors*
The complete ERIC vocabulary of descriptors is listed alphabetically, hierarchically, and in a rotated display.
Paperback, updated anually, $7.95
Available from Macmillan, Inc.

(¶ 1.119)

2. *Descriptor and Identifier Usage Report*
Lists each ERIC descriptor and identifier and the number of times it has been used to date
1966-77 Accumulated totals, $3.50
Available from the ERIC Clearinghouse on Science, Mathematics, and Environmental Education

(¶ 1.120)

3. *Institutional Source Directory*
An alphabetical Listing of the names of all institutions by which documents in the ERIC system have been indexed in the Corporate Source Name and Sponsoring Agency Name fields of the citation, together with the corresponding alphanumeric code which appears in either the Corporate Source or Sponsoring Agency Code field of each record
Published semianually, $15 domestic, $18 foreign
Annual subscription, $25 domestic, $30 foreign.
Available from the ERIC Processing and Reference Facility.

(¶ 1.121)

4. *Directory of ERIC Microfiche Collections*
A listing of the over 600 libraries and information centers that have a standing order to the ERIC microfiche collection. The directory gives collection scope/size, equipment and services and accessibility data, as well as the name and phone number of a contact person at each collection. Many of these facilities are open to the general public, and will duplicate microfiche and/or hardcopy.
Published annually; free on request
Available from the ERIC Processing and Reference Facility.

(¶ 1.122)

5. *Survey of ERIC Data Base Search Services*
A listing of the sites which provide search services on a regular basis—irrespective of whether the service is available only to a circumscribed community or to all users without restriction.
Published irregularly; free on request
Available from the ERIC Processing and Reference Facility.

(¶ 1.123)

6. *Interchange Newsletter*
The newsletter used by the ERIC Processing and Reference Facility to communicate with ERIC users about changes in the system, search tips, and periodically held ERIC users meetings.
Published irregularly; free of charge to interested organizations.

DOCUMENT RETRIEVAL

(¶ 1.124)

Copies of non-copyrighted documents (and any others for which reproduction permission has been obtained) announced in the RIE subfile of the ERIC database are available from the ERIC Document Reproduction Service (EDRS), P.O. Box 190, Arlington, Virginia 22210.

(¶ 1.126)

Copies may also be available from many of the facilities listed in the *Directory of ERIC Collections*, described in the Search Aids section.

(¶ 1.127)

INDEX